Глава 4.

Математическое приложение

В данной главе приводятся необходимые сведения из курсов функционального анализа и теории вероятностей, а также справочные сведения, используемые для вычисления специальных интегралов.

4.1. Необходимые сведения из функционального анализа

4.1.1. Алгебры и $\boldsymbol{\sigma}$-алгебры множеств

Определение 4.1.1. *Система $\mathcal{A}$ подмножеств некоторого множества X называется* ***алгеброй****, если*

1) $\varnothing, X \in \mathcal{A}$;

2) $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A},\ A \cap B \in \mathcal{A}$;

3) $A \in \mathcal{A} \implies X \setminus A = \overline{A} \in \mathcal{A}$.

Свойство 2 выполнено для любого конечного набора подмножеств, т. е. если $A_k \in \mathcal{A}$, $k = 1, \ldots, n$, то $\bigcup\limits_{k=1}^{n} A_k \in \mathcal{A}$, $\bigcap\limits_{k=1}^{n} A_k \in \mathcal{A}$.

Определение 4.1.2. *Алгебра $\mathcal{A}$ называется* ***$\boldsymbol{\sigma}$-алгеброй****, если*

$$A_n \in \mathcal{A},\ n = 1, 2, \ldots \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A},\ \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Определение 4.1.3. *Множество X вместе с некоторой σ-алгеброй его подмножеств $\mathcal{A}$ называется* ***измеримым пространством*** *и обозначается $\{X, \mathcal{A}\}$.*

На одном и том же множестве X могут быть заданы различные σ-алгебры его подмножеств. Примерами σ-алгебр являются системы множеств

$$\mathcal{A}_0 = \{\varnothing, X\}, \qquad \mathcal{A}^0 = \{A : A \subseteq X\}.$$

При этом $\mathcal{A}_0$ — самая "бедная" σ-алгебра, называемая ***тривиальной***, а $\mathcal{A}^0$ — самая "богатая" σ-алгебра, состоящая ***из всех подмножеств*** X.

Теорема 4.1.1. *Пусть на множестве X задана некоторая система его подмножеств $\mathcal{D}$. Тогда существует наименьшая σ-алгебра, обозначаемая $\sigma(\mathcal{D})$, содержащая все множества из $\mathcal{D}$.*

Система $\sigma(\mathcal{D})$ является наименьшей в том смысле, что если $\mathcal{A}$ — любая σ-алгебра подмножеств X, содержащая систему $\mathcal{D}$, то $\sigma(\mathcal{D}) \subseteq \mathcal{A}$.

Замечания: 1. Систему множеств $\sigma(\mathcal{D})$ называют ***σ-алгеброй, порожденной системой множеств*** $\mathcal{D}$.

2. Если $\{\mathcal{A}_\alpha\}$ — произвольное семейство σ-алгебр на X, то $\mathcal{A} = \bigcap\limits_\alpha \mathcal{A}_\alpha$ также является σ-алгеброй. Очевидно, что $\sigma(\mathcal{D}) = \bigcap\limits_\alpha \mathcal{A}_\alpha$, где $\{\mathcal{A}_\alpha\}$ — семейство всех σ-алгебр, содержащих систему множеств $\mathcal{D}$.

4.1.2. Меры (определения и свойства)

Определение 4.1.4. *Пусть на множестве X задана некоторая алгебра его подмножеств $\mathcal{A}$. Функция $\mu(A)$, определенная на множествах $A \in \mathcal{A}$, называется **мерой, заданной на** $\mathcal{A}$, если*

1) $\mu(A) \geq 0$ *для всех* $A \in \mathcal{A}$;

2) *для любого счетного набора попарно непересекающихся множеств $A_n \in \mathcal{A}$, $n = 1, 2, \ldots$, (т. е. $A_i \cap A_j$, $i \neq j$) таких, что $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{A}$ выполнено свойство **счетной аддитивности (σ-аддитивности)**:*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

*Мера μ называется **конечной**, если дополнительно выполнено условие:*

3) $\mu(X) < \infty$.

Мера μ, заданная на алгебре $\mathcal{A}$, обладает следующими свойствами:

1. $\mu(\varnothing) = 0$.

2. $\mu(A) \le \mu(B)$ для $A, B \in \mathcal{A}$ таких, что $A \subseteq B$.

3. $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$ для всех $A, B \in \mathcal{A}$.

4. Если $A_n \in \mathcal{A}$, $n = 1, 2, \ldots$ — убывающая последовательность множеств, т. е. $A_1 \supseteq A_2 \supseteq \ldots$, такая, что $\mu(A_1) < \infty$ и $\bigcap\limits_{n=1}^{\infty} A_n \in \mathcal{A}$, то

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

5. Если $A_n \in \mathcal{A}$, $n = 1, 2, \ldots$ — возрастающая последовательность множеств, т. е. $A_1 \subseteq A_2 \subseteq \ldots$, такая, что $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{A}$, то

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

Замечание. Если на алгебре $\mathcal{A}$ задана функция множества μ, обладающая свойством ***аддитивности***:

$$A, B \in \mathcal{A},\ A \cap B = \varnothing \quad \Longrightarrow \quad \mu(A \cup B) = \mu(A) + \mu(B),$$

то для доказательства того, что μ является мерой на $\mathcal{A}$, достаточно проверить свойство 4 для случая, когда $\bigcap\limits_{n=1}^{\infty} A_n = \varnothing$, т. е. $\mu(A_n) \to 0$ при $n \to \infty$.

Следующий результат является одним из наиболее важных результатов теории меры.

Теорема 4.1.2 (Каратеодори). *Пусть X — некоторое множество, $\mathcal{A}$ — алгебра его подмножеств и μ_0 — мера на $\mathcal{A}$. Если существует последовательность множеств $X_n \in \mathcal{A}$ таких, что*

$$\mu_0(X_n) < \infty, \qquad X = \bigcup_n X_n, \tag{4.1.1}$$

то существует и притом единственная мера μ, определенная на $\sigma(\mathcal{A})$, являющаяся ***продолжением*** *μ_0, т. е. такая, что*

$$\mu(A) = \mu_0(A) \quad \forall A \in \mathcal{A}.$$

Замечание. Для конечной меры условие (4.1.1), очевидно, выполнено.

Определение 4.1.5. *Множество X вместе с мерой μ, определенной на некоторой σ-алгебре $\mathcal{A}$, называется* ***пространством с мерой*** *и обозначается $\{X, \mathcal{A}, \mu\}$.*

Всякую σ-алгебру $\mathcal{A}$ можно пополнить множествами вида $A \cup N$, где $A \in \mathcal{A}$, а $N \subset A_0$ для некоторого $A_0 \in \mathcal{A}$, имеющего нулевую меру: $\mu(A_0) = 0$. Нетрудно проверить, что система множеств $\widetilde{\mathcal{A}}$, содержащая множества указанного вида, также является σ-алгеброй. Пространство с мерой $\{X, \widetilde{\mathcal{A}}, \mu\}$ называется ***полным***.

Определение 4.1.6. *Пусть $\{X, \mathcal{A}, \mu\}$ — полное пространство с мерой. Если некоторое свойство $\mathcal{P}$ выполняется для всех $x \in X_0 \subseteq X$, где $\mu(X \setminus X_0) = 0$, то мы будем говорить, что* ***свойство $\mathcal{P}$ выполнено почти всюду (по мере μ)***.

4.1.3. Способы задания мер

Следующие примеры измеримых пространств являются наиболее важными для теории вероятностей и случайных процессов.

Дискретное измеримое пространство

Пусть множество $X = \{x_1, x_2, \ldots\}$ не более чем счетно, а $\mathcal{A}$ — σ-алгебра всех подмножеств X. Всякая мера μ на ***дискретном измеримом пространстве*** $\{X, \mathcal{A}\}$ задается числами $\mu_n = \mu(\{x_n\}) \geq 0$:

$$\mu(A) = \sum_{n\colon x_n \in A} \mu_n,$$

где $A \in \mathcal{A}$ — любое подмножество X. Мера μ конечна, если

$$\mu(X) = \mu\left(\bigcup_{n=1}^{\infty} \{x_n\}\right) = \sum_{n=1}^{\infty} \mu(\{x_n\}) = \sum_{n=1}^{\infty} \mu_n < \infty.$$

Измеримое пространство $\{\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\}$

Пусть $X = \mathbb{R}^1$ — действительная прямая, а $\langle a, b\rangle$ — промежуток, т. е. одно из множеств вида

$$(a, b], \quad [a, b), \quad (a, b), \quad [a, b],$$

где $-\infty \le a \le b \le \infty$. Обозначим через $\mathcal{A}$ систему подмножеств $A \subseteq \mathbb{R}^1$, состоящих из конечных объединений непересекающихся промежутков:

$$A = \bigcup_{i=1}^{n} \langle a_i, b_i \rangle\,, \quad n < \infty. \tag{4.1.2}$$

Очевидно, система $\mathcal{A}$ образует алгебру, но не является σ-алгеброй.

Определение 4.1.7. *σ-алгебра, порожденная системой $\mathcal{A}$, обозначается $\mathcal{B}(\mathbb{R}^1)$ и называется **борелевской σ-алгеброй** множеств действительной прямой, а ее элементы — **борелевскими множествами**.*

Аналогично определяется измеримое пространство $\{[a,b], \mathcal{B}([a,b])\}$, где $\mathcal{B}([a,b])$ состоит из множеств $B \subseteq [a,b]$ таких, что $B \in \mathcal{B}(\mathbb{R}^1)$. Система $\mathcal{B}([a,b])$ называется ***борелевской σ-алгеброй*** отрезка $[a,b]$.

Определение 4.1.8. ***Мерой Лебега на*** $\mathbb{R}^1$ *называется мера λ, определенная на $\{\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\}$, такая, что*

$$\lambda(\langle a, b\rangle) = b - a.$$

для всех $-\infty \le a \le b \le \infty$.

Из теоремы 4.1.2 следует, что мера Лебега на $\mathbb{R}^1$ существует и единственна. Отметим, что мера Лебега на $\mathbb{R}^1$ не является конечной, так как $\lambda(\mathbb{R}^1) = \infty$.

Конечная мера на борелевской σ-алгебре прямой или отрезка может быть задана с помощью ***функции распределения***, которая определяется следующим образом.

Определение 4.1.9. *Функция $F(x)$, $x \in \mathbb{R}^1$ называется **функцией распределения** на $\mathbb{R}^1$, если она обладает следующими свойствами:*

1) *$F(x)$ — неубывающая функция;*

2) *$F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0$, $F(+\infty) = \lim\limits_{x\to+\infty} F(x) < \infty$;*

3) *$F(x)$ непрерывна справа, т. е. $\lim\limits_{h\to+0} F(x+h) = F(x)$ для всех $x \in \mathbb{R}^1$.*

Замечание. Для функции распределения на отрезке $[a,b]$ свойство 2 принимает вид

$$F(a) = \lim_{x\to a} F(x) = 0, \qquad F(b) = \lim_{x\to b} F(x) < \infty.$$

Теперь с использованием функции распределения $F(x)$ зададим функцию множества μ_0 на промежутках $\langle a, b\rangle$:

$$\mu_0((a,b]) = F(b) - F(a), \qquad \mu_0([a,b]) = F(b) - F(a-),$$
$$\mu_0((a,b)) = F(b-) - F(a), \qquad \mu_0([a,b)) = F(b-) - F(a-),$$

где $F(x-) = \lim\limits_{h\to+0} F(x-h)$. Далее, если $A = \bigcup\limits_{i=1}^{n} \langle a_i, b_i\rangle$, $n < \infty$, где промежутки $\langle a_i, b_i\rangle$ попарно не пересекаются, то положим

$$\mu_0\left(\bigcup_{i=1}^{n} \langle a_i, b_i\rangle\right) = \sum_{i=1}^{n} \mu_0(\langle a_i, b_i\rangle).$$

Тем самым мы задали μ_0 на системе $\mathcal{A}$, состоящей из множеств вида (4.1.2). В силу свойств функции распределения $F(x)$ (см. определение 4.1.9) μ_0 является конечной мерой на алгебре $\mathcal{A}$, поэтому по теореме 4.1.2 может быть продолжена и притом единственным образом до меры μ, заданной на измеримом пространстве $\{\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\}$.

З а м е ч а н и я: 1. С каждой конечной мерой μ, заданной на $\{\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\}$, можно связать функцию

$$F_\mu(x) = \mu((-\infty, x]), \quad x \in \mathbb{R}^1.$$

Из свойств меры (см. разд. 4.1.2) вытекает, что F_μ является функцией распределения, причем $F_\mu(+\infty) = \lim\limits_{x\to+\infty} F_\mu(x) = \mu(\mathbb{R}^1)$.

2. Между функциями распределения и конечными мерами на $\{\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\}$ существует взаимно однозначное соответствие, т. е. всякой конечной мере μ соответствует функция распределения $F_\mu(x)$, и наоборот, для всякой функции распределения $F(x)$ на $\mathbb{R}^1$ существует конечная мера μ такая, что $F_\mu \equiv F$.

Измеримое пространство $\{\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\}$

Пространство $\mathbb{R}^n$ есть прямое произведение n экземпляров прямых $\mathbb{R}^1$, т. е. $\mathbb{R}^n = \mathbb{R}^1 \times \ldots \times \mathbb{R}^1$ — множество упорядоченных наборов $x = (x_1, \ldots, x_n)^*$. Определим на этом пространстве систему подмножеств $\mathcal{A}^n$, образованную множествами

$$A = A_1 \times \ldots \times A_n = \prod_{k=1}^{n} A_k, \quad A_k \in \mathcal{A},$$

где $\mathcal{A}$ — алгебра подмножеств прямой вида (4.1.2). Нетрудно показать, что $\mathcal{A}^n$ образует алгебру.

Определение 4.1.10. *σ-алгебра, порожденная системой $\mathcal{A}^n$, обозначается $\mathcal{B}(\mathbb{R}^n)$ и называется **борелевской σ-алгеброй** множеств $\mathbb{R}^n$, а ее элементы — **борелевскими множествами**.*

Определение 4.1.11. ***Мерой Лебега на** $\mathbb{R}^n$ называется мера λ^n, определенная на $\{\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\}$, такая, что*

$$\lambda^n\left(\prod_{i=1}^{n} \langle a_i, b_i \rangle\right) = \prod_{i=1}^{n}(b_i - a_i)$$

для всех $-\infty \leq a_i \leq b_i \leq \infty$.

Из теоремы 4.1.2 следует, что мера Лебега на $\mathbb{R}^n$ существует и единственна. Отметим, что мера Лебега на $\mathbb{R}^n$ не является конечной, так как $\lambda^n(\mathbb{R}^n) = \infty$.

Конечная мера на борелевской σ-алгебре $\mathbb{R}^n$ может быть задана с помощью ***n*-мерной функции распределения**, которая определяется следующим образом.

Определение 4.1.12. *Функция $F(x_1, \ldots, x_n)$, $x_1, \ldots, x_n \in \mathbb{R}^1$ называется **n-мерной функцией распределения**, если она обладает следующими свойствами:*

1) *$F(x_1, \ldots, x_n)$ монотонна в следующем смысле:*

$$\Delta_1 \ldots \Delta_n F(x_1, \ldots, x_n) \geq 0;$$

где Δ_i — оператор конечной разности по переменной x_i:

$$\Delta_i F = F(x_1, \ldots, x_{i-1}, x_i + h_i, x_{i+1}, \ldots, x_n) - F(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n),$$

а $h_1 \geq 0, \ldots, h_n \geq 0$ произвольны;

2) *если хотя бы одна из переменных $x_i \to -\infty$, то*

$$F(x_1, \ldots, x_n) \to 0;$$

если все переменные $x_i \to +\infty$, то

$$F(x_1, \ldots, x_n) \to F(+\infty, \ldots, +\infty) < \infty;$$

3) *$F(x_1, \ldots, x_n)$ непрерывна справа по переменным x_i.*

Из теоремы 4.1.2 и свойств n-мерной функции распределения $F(x_1, \ldots, x_n)$ (см. определение 4.1.12) следует, что на измеримом пространстве $\{\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\}$ существует однозначно определенная конечная мера μ такая, что $F_\mu \equiv F$, где

$$F_\mu(x_1, \ldots, x_n) = \mu\left(\prod_{i=1}^{n}(-\infty, x_i]\right), \quad x_1, \ldots, x_n \in \mathbb{R}^1. \qquad (4.1.3)$$

Верно и обратное, если μ — конечная мера на $\{\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\}$, то функция $F_\mu(x_1, \ldots, x_n)$, определяемая соотношением (4.1.3), является n-мерной функцией распределения. Тем самым между n-мерными функциями распределения и конечными мерами на $\{\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\}$ существует взаимно однозначное соответствие. В этом случае $F_\mu(+\infty, \ldots, +\infty) = \mu(\mathbb{R}^n)$.

4.1.4. Измеримые функции

Пусть $\{X, \mathcal{A}\}$ — некоторое измеримое пространство.

Определение 4.1.13. *Вещественная функция $f(x)$, $x \in X$ называется **$\mathcal{A}$-измеримой**, если*

$$f^{-1}(B) \in \mathcal{A} \quad \forall B \in \mathcal{B}(\mathbb{R}^1), \qquad (4.1.4)$$

*где $f^{-1}(B) = \{x \in X : f(x) \in B\}$ — **прообраз** множества B.*

Тем самым измеримость функции означает то, что прообраз любого борелевского подмножества $\mathbb{R}^1$ является измеримым множеством в X.

Замечания: 1. Постоянная функция $f(x) \equiv \mathrm{const}$ очевидно является измеримой относительно любой σ-алгебры.

2. Индикаторная функция $I_A(x)$ множества A, определяемая как

$$I_A(x) = \begin{cases} 1, & \text{при} \quad x \in A, \\ 0, & \text{при} \quad x \notin A, \end{cases}$$

является измеримой в том и только том случае, когда множество A измеримо, т. е. $A \in \mathcal{A}$.

3. Условие $\mathcal{A}$-измеримости функции $f(x)$ тем ограничительнее, чем уже σ-алгебра $\mathcal{A}$, т. е. $\mathcal{A}$-измеримая $f(x)$ является $\mathcal{A}'$-измеримой, если $\mathcal{A} \subseteq \mathcal{A}'$.

Для проверки измеримости функции используется следующий результат.

Теорема 4.1.3. *Функция $f(x)$, заданная на измеримом пространстве $\{X, \mathcal{A}\}$, является $\mathcal{A}$-измеримой, если для всякого $c \in \mathbb{R}^1$ измеримыми являются множества*

$$\{x \in X : f(x) \leq c\}.$$

Функция $\varphi(y)$, $y \in \mathbb{R}^1$, заданная на действительной прямой, называется ***борелевской функцией***, если она $\mathcal{B}(\mathbb{R}^1)$-измерима. Примерами борелевских функций являются все кусочно-непрерывные функции.

Теорема 4.1.4. *Пусть $\{X, \mathcal{A}\}$ — измеримое пространство. Сложная функция $h(x) = \varphi(f(x))$, $x \in X$, является $\mathcal{A}$-измеримой, если функция $f(x)$, $x \in X$, $\mathcal{A}$-измерима, а функция $\varphi(y)$, $y \in \mathbb{R}^1$ является борелевской.*

Простые арифметические операции над конечным или счетным набором измеримых функций не выводят за рамки множества измеримых функций.

Теорема 4.1.5. *Пусть функции f_n, $n = 1, 2, \ldots$ определены на измеримом пространстве $\{X, \mathcal{A}\}$ и $\mathcal{A}$-измеримы. Тогда функции*

$$f_1(x) + f_2(x), \quad f_1(x)f_2(x), \quad 1/f_1(x) \ (\textit{при условии } f_1(x) \neq 0),$$

$$|f_1(x)|, \quad \max\{f_1(x), f_2(x)\}, \quad \min\{f_1(x), f_2(x)\}, \quad \sup_n f_n(x), \quad \inf_n f_n(x)$$

также являются измеримыми.

Из приведенного результата следует, что множество

$$A = \{x \in X : \exists \lim_{n\to\infty} f_n(x)\},$$

на котором существует предел последовательности измеримых функций $\{f_n(x)\}$, является измеримым, т. е. $A \in \mathcal{A}$.

Следующее определение описывает важный пример измеримой функции.

Определение 4.1.14. *Измеримая функция $f(x)$, определенная на измеримом пространстве $\{X, \mathcal{A}\}$, называется **простой**, если она принимает конечное число значений.*

Очевидно, что каждая простая измеримая функция $f(x)$ допускает представление

$$f(x) = \sum_{k=1}^{m} c_k I_{A_k}(x), \quad x \in X, \tag{4.1.5}$$

где $c_k \in \mathbb{R}^1$, $A_k \in \mathcal{A}$ и $A_i \cap A_j = \varnothing$, $i \neq j$, $m < \infty$.

Теорема 4.1.6. *Для всякой неотрицательной измеримой функции $f(x)$ существует неубывающая последовательность неотрицательных простых измеримых функций $\{f_n(x)\}$, сходящаяся к $f(x)$, т. е.*

$$0 \le f_n(x) \uparrow f(x) \quad \text{при} \quad n \to \infty \quad \text{для всех} \quad x \in X.$$

Результат, приведенный выше, является ключевым при построении интеграла Лебега.

Определение 4.1.15. *Пусть $f(x)$ — некоторая измеримая функция, определенная на измеримом пространстве $\{X, \mathcal{A}\}$. Система*

$$\mathcal{A}_f = \{f^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^1)\}$$

называется ***σ-алгеброй, порожденной функцией*** *$f(x)$.*

Нетрудно убедиться, что $\mathcal{A}_f$ действительно является σ-алгеброй, причем $\mathcal{A}_f \subseteq \mathcal{A}$. Класс функций, измеримых относительно σ-алгебры $\mathcal{A}_f$, имеет простое описание.

Теорема 4.1.7. *Функция $g(x)$, $x \in X$ является $\mathcal{A}_f$-измеримой тогда и только тогда, когда существует борелевская функция $\varphi(y)$, $y \in \mathbb{R}^1$ такая, что*

$$g(x) = \varphi(f(x)) \quad \forall x \in X.$$

Определение 4.1.16. *Две измеримые функции $f(x)$ и $g(x)$, заданные на $\{X, \mathcal{A}, \mu\}$, называются* ***эквивалентными****, если $f(x) = g(x)$ почти всюду по мере μ, т. е. $\mu\{x \in X : f(x) \ne g(x)\} = 0$.*

Определение 4.1.17. *Последовательность $\{f_n(x)\}$ измеримых функций, определенных на $\{X, \mathcal{A}, \mu\}$, называется* ***сходящейся почти всюду*** *к функции $f(x)$, если*

$$\mu\{x \in X : \lim_{n\to\infty} f_n(x) \ne f(x)\} = 0.$$

Теорема 4.1.8. *Если последовательность $\{f_n(x)\}$ измеримых функций сходится к $f(x)$ почти всюду, то $f(x)$ измерима.*

Определение 4.1.18. *Последовательность $\{f_n(x)\}$ измеримых функций, определенных на $\{X, \mathcal{A}, \mu\}$, называется* ***сходящейся по мере*** *μ к измеримой функции $f(x)$, если для любого $\varepsilon > 0$*

$$\lim_{n\to\infty} \mu\{x \in X : |f_n(x) - f(x)| > \varepsilon\} = 0.$$

Соотношение между этими двумя типами сходимости определяется следующими теоремами.

Теорема 4.1.9. *Если последовательность $\{f_n(x)\}$ сходится к $f(x)$ почти всюду относительно конечной меры μ, то $f_n(x)$ сходится к $f(x)$ по мере μ.*

Обратное утверждение, вообще говоря, не верно, тем не менее, справедлив следующий результат.

Теорема 4.1.10. *Если последовательность $\{f_n(x)\}$ сходится к $f(x)$ по мере, то существует подпоследовательность $\{f_{n_k}(x)\}$, сходящаяся к $f(x)$ почти всюду.*

4.1.5. Интеграл Лебега

Пусть $\{X, \mathcal{A}, \mu\}$ — полное пространство с мерой. Определим вначале интеграл Лебега от простой измеримой функции.

Определение 4.1.19. ***Интеграл Лебега от простой измеримой функции*** $f(x)$, *имеющей вид* (4.1.5), *определяется равенством*

$$\int\limits_X f(x)\,\mu(dx) = \sum_k c_k \mu(A_k). \tag{4.1.6}$$

Теперь распространим понятие интеграла Лебега на неотрицательную измеримую функцию $f(x)$. В силу теоремы 4.1.6 существует последовательность неотрицательных простых измеримых функций $\{f_n(x)\}$ таких, что

$$0 \le f_n(x) \uparrow f(x) \quad \text{при} \quad n \to \infty \quad \text{для всех} \quad x \in X. \tag{4.1.7}$$

Определение 4.1.20. ***Интегралом Лебега от неотрицательной измеримой функции*** $f(x)$ *называется величина*

$$\int\limits_X f(x)\,\mu(dx) = \lim_{n\to\infty} \int\limits_X f_n(x)\,\mu(dx). \tag{4.1.8}$$

Данное определение корректно, поскольку при заданной функции $f(x)$ предел (4.1.8) (конечный или бесконечный) существует для любой аппроксимирующей последовательности (4.1.7) и не зависит от ее выбора.

Для краткости будем обозначать

$$I(f) = \int\limits_X f(x)\,\mu(dx).$$

Пусть теперь $f(x)$ — произвольная измеримая функция. Обозначим

$$f^+(x) = \max\{f(x), 0\}, \qquad f^-(x) = \max\{-f(x), 0\}.$$

Тогда функции $f^+(x) \geq 0$, $f^-(x) \geq 0$ измеримы и $f(x) = f^+(x) - f^-(x)$.

Определение 4.1.21. *Если* $\min\{I(f^+), I(f^-)\} < \infty$, *то* ***интеграл Лебега от*** $f(x)$ ***существует*** *(или* ***определен****) и имеет вид*

$$I(f) = \int\limits_X f(x)\,\mu(dx) = I(f^+) - I(f^-).$$

Если $|I(f)| < \infty$, *то* $f(x)$ *называется* ***интегрируемой по мере*** μ *или* ***суммируемой****.*

Интеграл по множеству $A \in \mathcal{A}$ *определяется как*

$$\int\limits_A f(x)\,\mu(dx) = \int\limits_X f(x) I_A(x)\,\mu(dx),$$

а $f(x)$ *называется* ***интегрируемой на множестве*** A, *если интегрируема функция* $f(x)I_A(x)$.

Замечание. В соответствии с определением 4.1.21 интеграл Лебега от неотрицательной функции $f(x)$ определен всегда, при этом $0 \leq I(f) \leq \infty$.

Перечислим свойства интеграла Лебега, непосредственно вытекающие из его определения.

1. $\int\limits_X I_A(x)\,\mu(dx) = \mu(A)$ для всякого $A \in \mathcal{A}$.

2. Для всех α, $\beta \in \mathbb{R}^1$

$$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g),$$

причем если правая часть имеет смысл, то интеграл в левой части определен.

3. Если $f(x)$, $g(x)$ интегрируемы и $f(x) \leq g(x)$, то $I(f) \leq I(g)$.

4. Если $\mu(A) = 0$, то $\int\limits_A f(x)\,\mu(dx) = 0$.

5. Если $f(x) = g(x)$ почти всюду по мере μ, то $I(f) = I(g)$, причем оба интеграла существуют или не существуют одновременно.

6. Если $h(x) \geq 0$ интегрируема и $|f(x)| \leq h(x)$ почти всюду по мере μ, то $f(x)$ интегрируема.

7. Интегралы $I(f)$ и $I(|f|)$ существуют или не существуют одновременно, при этом интегрируемость функции равносильна интегрируемости ее модуля, т. е.

$$|I(f)| < \infty \iff I(|f|) < \infty,$$

кроме того, $|I(f)| \leq I(|f|)$.

8. Если $f(x) \geq 0$ и $I(f) = 0$, то $f(x) = 0$ почти всюду. В частности, если $I(|f|) = 0$, то $f(x) = 0$ почти всюду.

9. Если функция $f(x)$ интегрируема на A, то она интегрируема на любом измеримом множестве $A' \subseteq A$.

10. Если $f(x) = 0$ почти всюду, то $I(f) = 0$.

11. (Неравенство Чебышева). Если $c > 0$, то

$$\mu\{x \in X : |f(x)| \geq c\} \leq \frac{1}{c} \int\limits_X |f(x)|\, \mu(dx).$$

12. (Неравенство Минковского). Если $p \geq 1$, то

$$\left(\int\limits_X |f(x) + g(x)|^p \, \mu(dx) \right)^{1/p} \leq \left(\int\limits_X |f(x)|^p \, \mu(dx) \right)^{1/p} + \left(\int\limits_X |g(x)|^p \, \mu(dx) \right)^{1/p}.$$

13. (Неравенство Иенсена). Если $\varphi(y)$, $y \in \mathbb{R}^1$ выпукла и $\mu(X) = 1$, то

$$\int\limits_X \varphi(f(x))\, \mu(dx) \leq \varphi\left(\int\limits_X f(x)\, \mu(dx) \right).$$

14. (Неравенство Гельдера). Если $1 < p$, $q < \infty$ и $\dfrac{1}{p} + \dfrac{1}{q} = 1$, то

$$\left| \int\limits_X f(x) g(x)\, \mu(dx) \right| \leq \left(\int\limits_X |f(x)|^p \, \mu(dx) \right)^{1/p} \cdot \left(\int\limits_X |g(x)|^q \, \mu(dx) \right)^{1/q},$$

причем интеграл в левой части определен, если интегралы в правой части конечны. При $p = q = 2$ неравенство Гельдера называется ***неравенством Коши–Буняковского***:

$$\left| \int\limits_X f(x)g(x)\mu(dx) \right|^2 \leq \int\limits_X |f(x)|^2\mu(dx) \int\limits_X |g(x)|^2\mu(dx).$$

Теорема 4.1.11 (σ-аддитивность интеграла Лебега). *Пусть функция $f(x)$ интегрируема. Если $X = \bigcup\limits_{n=1}^{\infty} A_n$, где $A_i \cap A_j = \varnothing$ при $i \neq j$, то*

$$\int\limits_X f(x)\,\mu(dx) = \sum_{n=1}^{\infty} \int\limits_{A_n} f(x)\,\mu(dx),$$

причем в правой части интегралы конечны, а ряд сходится абсолютно.

Из теоремы 4.1.11 следует, что для любой неотрицательной измеримой функции $f(x) \geq 0$ функция множества

$$\nu(A) = \int\limits_A f(x)\,\mu(dx), \quad A \in \mathcal{A} \tag{4.1.9}$$

является мерой на измеримом пространстве $\{X, \mathcal{A}\}$.

Теорема 4.1.12 (абсолютная непрерывность интеграла Лебега). *Пусть $f(x)$ — интегрируемая функция, тогда для любого $\varepsilon > 0$ существует такое $\delta > 0$, что $\left| \int\limits_A f(x)\,\mu(dx) \right| < \varepsilon$ для всякого измеримого множества $A \in \mathcal{A}$ такого, что $\mu(A) < \delta$.*

Следующая теорема утверждает, что всякая мера ν, абсолютно-непрерывная относительно меры μ, допускает представление (4.1.9).

Определение 4.1.22. *Мера ν называется **абсолютно-непрерывной относительно меры** μ (сокращенно $\nu \ll \mu$), если для любого $A \in \mathcal{A}$ такого, что $\mu(A) = 0$ следует, что и $\nu(A) = 0$.*

Теорема 4.1.13 (Радон–Никодим). *Если $\nu \ll \mu$, то существует измеримая неотрицательная функция $\rho(x)$ такая, что*

$$\nu(A) = \int\limits_A \rho(x)\,\mu(dx) \quad \forall A \in \mathcal{A}.$$

Функция $\rho(x)$ называется ***производной Радона–Никодима меры*** ν ***по мере*** μ и обозначается

$$\rho(x) = \frac{d\nu}{d\mu}(x).$$

Следующий результат дает правило замены меры в интеграле Лебега.

Теорема 4.1.14. *В условиях теоремы 4.1.13 для любой измеримой функции $g(x)$ имеет место равенство*

$$\int\limits_X g(x)\,\nu(dx) = \int\limits_X g(x)\,\rho(x)\,\mu(dx),$$

где интегралы в левой и правой частях существуют или не существуют одновременно.

Тем самым интегрируемость функции $g(x)$ по мере ν равносильна интегрируемости функции $g(x)\,\rho(x)$ по мере μ.

Теперь рассмотрим вопрос о замене переменной под знаком интеграла Лебега.

Определение 4.1.23. *Пусть $f(x)$ — измеримая функция, определенная на пространстве с мерой $\{X, \mathcal{A}, \mu\}$. Тогда мера μ_f на $\{\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\}$, задаваемая равенством*

$$\mu_f(B) = \mu(f^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}^1),$$

называется ***мерой, порожденной функцией*** f.

Нетрудно проверить, что функция множества μ_f действительно является мерой на борелевской σ-алгебре $\mathcal{B}(\mathbb{R}^1)$.

Теорема 4.1.15 (формула замены переменной в интеграле Лебега). *Пусть функция $f(x)$, $x \in X$ измерима, а функция $\varphi(y)$, $y \in \mathbb{R}^1$ является борелевской. Тогда справедливо равенство*

$$\int\limits_X \varphi(f(x))\,\mu(dx) = \int\limits_{\mathbb{R}^1} \varphi(y)\,\mu_f(dy),$$

где интегралы в левой и правой частях существуют или не существуют одновременно.

Предельный переход под знаком интеграла Лебега

Следующие результаты показывают, в каких случаях можно переходить к пределу под знаком интеграла Лебега.

Теорема 4.1.16 (А. Лебег). *Пусть последовательность* $\{f_n(x)\}$ *на* X *сходится почти всюду к* $f(x)$, *причем для всех* n *почти всюду выполняется неравенство* $|f_n(x)| \leq \varphi(x)$, *где функция* $\varphi(x)$ *интегрируема. Тогда предельная функция* $f(x)$ *также является интегрируемой и*

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

Теорема 4.1.17 (Б. Леви). *Пусть* $\{f_n(x)\}$ — *неубывающая последовательность функций, т. е.*

$$f_1(x) \leq f_2(x) \leq \ldots \leq f_n(x) \leq \ldots,$$

где $f_n(x)$ *интегрируемы, а их интегралы ограничены в совокупности, т. е.*

$$\exists K < \infty : \quad \int\limits_X f_n(x)\,\mu(dx) \leq K.$$

Тогда почти всюду существует конечный предел $f(x) = \lim\limits_{n\to\infty} f_n(x)$ *такой, что функция* $f(x)$ *интегрируема и*

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

Теорема 4.1.18 (П. Фату). *Если последовательность неотрицательных функций* $\{f_n(x)\}$ *сходится почти всюду к* $f(x)$ *и* $\int\limits_X f_n(x)\,\mu(dx) \leq K$ *для всех* n, *то* $f(x)$ *интегрируема и* $\int\limits_X f(x)\,\mu(dx) \leq K$.

Интеграл Лебега на прямой

Следующее утверждение показывает связь между интегралом Римана и интегралом Лебега.

Теорема 4.1.19. *Если существует интеграл Римана*

$$I_R(f) = \int_a^b f(x)\,dx,$$

то $f(x)$ *интегрируема по мере Лебега* λ *на* $[a,b]$ *и*

$$I_L(f) = \int_{[a,b]} f(x)\,\lambda(dx) = I_R(f).$$

Заметим, что обратное утверждение в общем случае не верно.

Введем еще одно понятие интеграла, основанное на понятии интеграла Лебега.

Определение 4.1.24. *Пусть* $f(x)$, $x \in \mathbb{R}^1$ *— борелевская функция, а* $F(x)$ *— функция распределения на* $\mathbb{R}^1$ *(см. определение 4.1.9).* ***Интеграл Стилтьеса от*** $f(x)$ ***по функции распределения*** $F(x)$ *определяется следующим образом:*

$$\int_{-\infty}^{\infty} f(x)\,dF(x) = \int_{\mathbb{R}^1} f(x)\,\mu(dx),$$

где интеграл в правой части понимается как интеграл Лебега по мере μ*, имеющей функцию распределения* $F(x)$.

В большинстве практически важных случаев функция распределения $F(x)$ может быть представлена в виде

$$F(x) = F^a(x) + F^d(x), \quad F^a(x) = \int_{-\infty}^{x} p(y)\,\lambda(dy), \quad F^d(x) = \sum_{k\colon x_k \le x} p_k,$$

где функция $p(y) \ge 0$ интегрируема по мере Лебега на $\mathbb{R}^1$, а множество точек $\{x_k\}$ не более чем счетно, причем $p_k > 0$.

Составляющая $F^a(x)$ называется ***функцией абсолютно-непрерывного распределения*** или просто ***абсолютно-непрерывной функцией***. В этом случае для почти всех $x \in \mathbb{R}^1$ (по мере Лебега) справедливо равенство $\dfrac{d}{dx}F(x) = \dfrac{d}{dx}F^a(x) = p(x)$, поэтому $p(x)$ называется ***функцией плотности распределения***. Заметим, что если $\mu_a(dx)$ — мера, имеющая функцию распределения $F^a(x)$, то плотность $p(x)$ есть производная Радона–Никодима меры μ_a по мере Лебега λ.

Составляющая $F^d(x)$ называется ***функцией дискретного распределения***, при этом $p_k = F(x_k) - F(x_k-)$ есть величина скачка функции $F(x)$ в точке ее разрыва $x_k \in \mathbb{R}^1$. Если число скачков конечно, то $F^d(x)$ кусочно-постоянна.

Если $f(x)$ — борелевская функция, то справедливо следующее правило вычисление интеграла Стилтьеса:

$$\int\limits_{-\infty}^{\infty} f(x)\, dF(x) = \int\limits_{\mathbb{R}^1} f(x)\, p(x)\, \lambda(dx) + \sum_k f(x_k)\, p_k,$$

причем интеграл Стилтьеса конечен, если функция $f(x)\, p(x)$ интегрируема по мере Лебега, а ряд $\sum\limits_k f(x_k)\, p_k$ сходится абсолютно.

В заключение рассмотрим пример меры на прямой, чрезвычайно важный для приложения. Мера $\delta_{x_0}(B)$, $B \in \mathcal{B}(\mathbb{R}^1)$, $x_0 \in \mathbb{R}^1$, определяемая равенством

$$\delta_{x_0}(B) = \begin{cases} 1, & \text{если} \quad x_0 \in B, \\ 0, & \text{если} \quad x_0 \notin B, \end{cases}$$

называется ***мерой Дирака, сосредоточенной в точке*** x_0. Тогда интеграл по мере Дирака принимает вид

$$\int\limits_{\mathbb{R}^1} f(x)\, \delta_{x_0}(dx) = f(x_0).$$

Последний интеграл часто записывают в виде

$$\int\limits_{-\infty}^{\infty} f(x)\, \delta(x - x_0)\, dx = f(x_0),$$

где $\delta(x)$ называют ***$\boldsymbol{\delta}$-функцией Дирака***. При этом, поскольку функция

$$\mathbb{I}(x) = \begin{cases} 1, & \text{если} \quad x \geq 0, \\ 0, & \text{если} \quad x < 0 \end{cases}$$

является функцией распределения меры Дирака δ_0, считают, что выполнено равенство $\delta(x) = \dfrac{d\,\mathbb{I}(x)}{dx}$.

4.1.6. Гильбертово пространство

Сначала введем понятие линейного пространства.

Определение 4.1.25. *Непустое множество H называется **линейным пространством**, если на H определены две операции: **сложение элементов** $(+)$ и **умножение элементов на числа** $(\circ)$.*

Для любых элементов $x, y, z \in H$ и любых чисел α, β операции сложения и умножения должны обладать следующими свойствами:

1) $x + y = y + x$;

2) $(x + y) + z = x + (y + z)$;

3) *существует элемент* $0 \in H$ *такой, что* $0 + x = x$;

4) *существует элемент* $(-x) \in H$ *такой, что* $x + (-x) = 0$;

5) $\alpha \circ (x + y) = \alpha \circ x + \alpha \circ y$;

6) $(\alpha + \beta) \circ x = \alpha \circ x + \beta \circ y$;

7) $(\alpha\beta) \circ x = \alpha \circ (\beta \circ x)$;

8) $1 \circ x = x$.

*Если умножение определено на вещественные числа, то пространство называется **вещественным** или **действительным**, если умножение определено на комплексные числа, то пространство называется **комплексным**.*

Замечание. Во всяком линейном пространстве ***нулевой элемент*** *0* единственен, а ***противоположный элемент*** $(-x)$ однозначно определяется элементом $x \in H$. В дальнейшем мы будем опускать знак умножения $(\circ)$, а выражение $x + (-y)$ будем записывать в виде $x - y$.

Далее без ограничения общности будем считать, что H — комплексное линейное пространство, а множество комплексных чисел обозначать $\mathbb{C}$.

Определение 4.1.26. ***Подпространством*** *M линейного пространства H называется подмножество, замкнутое относительно операций сложения и умножения, т. е.*

$$\alpha, \beta \in \mathbb{C}, \ x, y \in M \implies \alpha x + \beta y \in M.$$

З а м е ч а н и я: 1. Определение 4.1.26 означает, что подпространство M вместе с каждым конечным набором своих элементов $\{x_1, \ldots, x_n\}$, содержит любую их ***линейную комбинацию*** $\alpha_1 x_1 + \ldots + \alpha_n x_n$, где $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. В частности, всегда верно $0 \in M$.

2. $M_0 = \{0\}$ называется ***нулевым подпространством***.

3. Пересечение $\bigcap\limits_{\alpha} M_\alpha$ произвольного набора подпространств $\{M_\alpha\}$ также является подпространством.

О п р е д е л е н и е 4.1.27. *__Линейной оболочкой__ $\mathcal{L}(N)$ некоторого множества $N \subseteq H$ называется множество всех возможных линейных комбинаций элементов из N.*

Из определения 4.1.27 следует, что $\mathcal{L}(N)$ является подпространством, причем $\mathcal{L}(N) = \bigcap\limits_{\alpha} M_\alpha$, где $\{M_\alpha\}$ — семейство всех подпространств, содержащих множество N.

О п р е д е л е н и е 4.1.28. *Элементы $\{x_1, \ldots, x_n\}$ линейного пространства H называются __линейно независимыми__, если*

$$\alpha_1, \ldots, \alpha_n \in \mathbb{C}, \ \alpha_1 x_1 + \ldots + \alpha_n x_n = 0 \implies \alpha_1 = \ldots = \alpha_n = 0.$$

Если для линейного пространства H существует набор линейно независимых элементов $\{x_1, \ldots, x_n\} \subset H$ такой, что $H = \mathcal{L}\{x_1, \ldots, x_n\}$, то набор $\{x_1, \ldots, x_n\}$ называют ***базисом*** пространства H, а число n — его ***размерностью*** (сокращенно $n = \dim H$). В этом случае каждый элемент $x \in H$ имеет представление $x = \alpha_1 x_1 + \ldots + \alpha_n x_n$, причем коэффициенты $\{\alpha_k\}$ определены однозначно.

О п р е д е л е н и е 4.1.29. *Функция (x, y), определенная на линейном пространстве H и принимающая числовые значения, называется __скалярным произведением__, если для любых элементов $x, y, z \in H$ и любых чисел α, β справедливы следующие свойства:*

1) $(x, x) \geq 0$ *и* $(x, x) = 0 \implies x = 0$;

2) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$;

3) $(y, x) = \overline{(x, y)}$.

О п р е д е л е н и е 4.1.30. *Пусть на H определено скалярное произведение. __Нормой__ элемента $x \in H$ будем называть число*

$$\|x\| = (x, x)^{1/2}. \tag{4.1.10}$$

В этом случае говорят, что норма ***порождена*** *скалярным произведением.*

Непосредственно из определений вытекают следующие свойства скалярного произведения и порожденной им нормы:

1) $\|x\| \geq 0$; из $\|x\| = 0$ следует $x = 0$;

2) $\|\alpha x\| = |\alpha| \|x\|$;

3) $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника);

4) $|(x, y)| \leq \|x\| \|y\|$ (неравенство Коши–Буняковского), причем равенство достигается только, если x, y линейно зависимы.

С помощью нормы можно ввести понятие ***сходимости последовательности*** элементов $\{x_n\}$ пространства H.

Определение 4.1.31. *Последовательность* $\{x_n\}$ ***сходится*** *к* $x \in H$ *при* $n \to \infty$ *(сокращенно* $x_n \to x$*), если* $\|x_n - x\| \to 0$.

Скалярное произведение является ***непрерывным*** в следующем смысле:

$$x_n \to x,\ y_m \to y \implies (x_n, y_m) \to (x, y).$$

Определение 4.1.32. *Последовательность* $\{x_n\}$ *называется* ***фундаментальной***, *если*

$$\|x_n - x_m\| \to 0 \quad \textit{при} \quad n, m \to \infty.$$

Определение 4.1.33. *Линейное пространство с нормой* (4.1.10) *называется* ***полным***, *если любая фундаментальная последовательность его элементов сходится к некоторому элементу этого пространства.*

Определение 4.1.34. *Линейное пространство со скалярным произведением и нормой* (4.1.10) *называется* ***гильбертовым пространством***, *если оно полно.*

Определение 4.1.35. *Подпространство* M *гильбертова пространства* H *называется* ***всюду плотным***, *если для всякого элемента* $x \in H$ *существует последовательность* $\{x_n\}$ *такая, что* $x_n \in M$ *и* $x_n \to x$.

Замечание. Если пространство H не является полным, то оно всегда может быть ***пополнено***, т. е. существует полное пространство $\widetilde{H}$ такое, что $H \subset \widetilde{H}$ и H всюду плотно в $\widetilde{H}$.

Простой критерий сходимости последовательности элементов гильбертова пространства приводится в следующей теореме.

Теорема 4.1.20. *Если H — гильбертово пространство, то последовательность $\{x_n\}$ сходится к некоторому элементу пространства H тогда и только тогда, когда существует предел* $\lim\limits_{n,m\to\infty}(x_n, x_m)$.

Определение 4.1.36. *Подпространство M гильбертова пространства H называется* ***замкнутым****, если из того, что $x_n \to x$, $n \to \infty$, где $x_n \in M$, следует $x \in M$.*

Таким образом, замкнутое подпространство гильбертова пространства само является гильбертовым пространством.

Рассмотрим важнейшие примеры гильбертовых пространств.

Конечномерное евклидово пространство $\mathbb{R}^n$

Его элементами являются упорядоченные наборы вещественных чисел (n-мерные векторы) $x = \{x_1, \ldots, x_n\}^*$, $x_1, \ldots, x_n \in \mathbb{R}^1$. Скалярное произведение и норма пространства $\mathbb{R}^n$ имеют вид

$$(x, y) = \sum_{k=1}^{n} x_k y_k, \qquad \|x\| = |x| = \left(\sum_{k=1}^{n} x_k^2\right)^{1/2},$$

где $y = \{y_1, \ldots, y_n\}^*$. Очевидно, что $\dim \mathbb{R}^n = n$.

Пространство l_2

Его элементами являются бесконечные последовательности комплексных чисел $x = \{x_1, x_2, \ldots\}$ таких, что

$$\|x\| = \left(\sum_{k=1}^{\infty} |x_k|^2\right)^{1/2} < \infty.$$

Скалярное произведение элементов $x, y \in l_2$ имеет вид

$$(x, y) = \sum_{k=1}^{\infty} x_k \overline{y_k}$$

и порождает норму $\|x\|$, определенную выше. Заметим, что пространство l_2 бесконечномерно.

Пространство $L_2\{X, \mathcal{A}, \mu\}$.

Пусть $\{X, \mathcal{A}, \mu\}$ — некоторое пространство с мерой. Тогда $L_2\{X, \mathcal{A}, \mu\}$ определяется как множество всех $\mathcal{A}$-измеримых комплексных функций $f(x)$, интегрируемых с квадратом по мере μ, т. е. таких, что

$$\|f\| = \left(\int\limits_X |f(x)|^2 \, \mu(dx) \right)^{1/2} < \infty. \tag{4.1.11}$$

Объединим в один класс эквивалентности все функции $\tilde{f}(x)$, эквивалентные (по мере μ) данной функции $f(x)$, и в дальнейшем функции $\tilde{f}(x)$ различать не будем. Тогда из $\|f\| = 0$ следует $f = 0$, поэтому выражение (4.1.11) определяет норму, порожденную скалярным произведением

$$(f, g) = \int\limits_X f(x) \, \overline{g(x)} \, \mu(dx), \quad f, g \in L_2\{X, \mathcal{A}, \mu\}. \tag{4.1.12}$$

Следовательно, $L_2\{X, \mathcal{A}, \mu\}$ является гильбертовым пространством. Отметим, что $L_2\{X, \mathcal{A}, \mu\}$ бесконечномерно, если мера μ не сосредоточена в конечном числе точек.

4.1.7. Ряды Фурье в гильбертовом пространстве

Пусть далее H — фиксированное гильбертово пространство.

Определение 4.1.37. *Два элемента $x, y \in H$ называются* ***ортогональными*** *(обозначается $x \perp y$), если $(x, y) = 0$.*

Определение 4.1.38. *Система $\{e_1, e_2, \ldots\}$ элементов пространства H называется* ***ортогональной****, если $e_m \perp e_n$ при $m \neq n$. Если дополнительно $\|e_n\| = 1$, $n = 1, 2, \ldots$, то система называется* ***ортонормальной****.*

Определение 4.1.39. *Ортонормальная система $\{e_1, e_2, \ldots\} \subset H$ называется* ***базисом гильбертова пространства*** *H, если для каждого $x \in H$ существует и притом единственное разложение*

$$x = \sum_{k=1}^{\infty} \alpha_k e_k, \tag{4.1.13}$$

где $\alpha_k \in \mathbb{C}$, а сходимость ряда понимается в следующем смысле:

$$\left\| x - \sum_{k=1}^{n} \alpha_k e_k \right\| \to 0 \quad \text{при} \quad n \to \infty.$$

Разложение (4.1.13) называется ***рядом Фурье*** для x, а коэффициенты α_k вычисляются по формулам

$$\alpha_k = (x, e_k), \quad k = 1, 2, \ldots$$

и называются ***коэффициентами Фурье***.

Из (4.1.13) также следует ***равенство Парсеваля***:

$$\|x\|^2 = \sum_{k=1}^{\infty} |\alpha_k|^2,$$

которое является обобщением теоремы Пифагора на бесконечномерный случай.

Теорема 4.1.21. *Пусть $\{e_n\}$ — ортонормальная система элементов гильбертова пространства H. Следующие утверждения равносильны:*

1) *$\{e_n\}$ — базис гильбертова пространства H;*

2) *линейная оболочка $\mathcal{L}(\{e_n\})$ всюду плотна в H;*

3) *если $x \perp e_n$ для всех n, то $x = 0$.*

В частности, система $\{e^{inx}/\sqrt{2\pi},\ n \in \mathbb{Z}\}$ является базисом пространства $L_2([-\pi, \pi], \mathcal{B}([-\pi, \pi]), \lambda)$, где λ — мера Лебега.

4.1.8. Ортогональное проектирование в гильбертовом пространстве

Определение 4.1.40. *Если $N \subseteq H$ — произвольное подмножество, то ортогональность $x \perp N$ означает, что $x \perp y$ для всех $y \in N$.*

Определение 4.1.41. *Если M — замкнутое подпространство, то его* ***ортогональное дополнение*** *$M^\perp$ определяется как*

$$M^\perp = \{x \in H : x \perp M\}.$$

В силу непрерывности скалярного произведения $M^\perp$ является замкнутым подпространством.

Теорема 4.1.22. *Пусть M — произвольное замкнутое подпространство гильбертова пространства H. Тогда любой элемент $x \in H$ имеет единственное разложение вида*

$$x = y_x + z_x,$$

где $y_x \in M$, $z_x \in M^\perp$.

Определение 4.1.42. *Элемент* y_x, *определенный в теореме* 4.1.22, *называется* ***ортогональной проекцией*** x ***на*** M *и обозначается* $\pi_M(x)$.

Перечислим основные свойства проекции $\pi_M(x)$:

1) $y_x = \pi_M(x) \in M$, $z_x = x - \pi_M(x) \in M^\perp$;

2) $x \in M$ тогда и только тогда, когда $\pi_M(x) = x$;
$x \in M^\perp$ тогда и только тогда, когда $\pi_M(x) = 0$;

3) $\|x - \pi_M(x)\| \le \|x - v\|$ для всех $v \in M$;

4) если $\{e_1, e_2, \ldots\}$ — базис гильбертова пространства H, а M является линейной оболочкой системы $\{e_1, \ldots, e_n\}$, т. е. $M = \mathcal{L}\{e_1, \ldots, e_n\}$, то

$$\pi_M(x) = \sum_{k=1}^{n} \alpha_k e_k,$$

где $\alpha_k = (x, e_k)$ — коэффициенты разложения x в ряд Фурье по базису $\{e_n\}$. Таким образом, $\pi_M(x)$ — n-я частичная сумма ряда Фурье для x.

Замечание. Свойство 3 означает, что $\pi_M(x)$ — наилучшая аппроксимация (оценка) элемента x элементами из подпространства M, а $x - \pi_M(x)$ представляет собой ошибку указанной аппроксимации. В частности, если $M = \mathcal{L}\{e_1, \ldots, e_n\}$, то $\pi_M(x)$ является наилучшей оценкой среди всех оценок, представимых в виде линейной комбинации элементов $\{e_1, \ldots, e_n\}$, т. е. является ***наилучшей линейной оценкой***.

Метод построения указанных аппроксимаций (с использованием проекции π_M) называют ***методом наименьших квадратов***.

4.2. Необходимые сведения из теории вероятностей

4.2.1. Случайные события и их вероятности

Определение 4.2.1. *Совокупность объектов* $\{\Omega, \mathcal{F}, \mathbf{P}\}$, *где*

Ω — *пространство* ***элементарных событий*** ω;

$\mathcal{F}$ — ***σ-алгебра подмножеств*** *пространства* Ω, *образующих систему* ***случайных событий***;

$\mathbf{P}$ — ***нормированная*** *(т. е.* $\mathbf{P}\{\Omega\} = 1$*)* ***мера*** *на* $\mathcal{F}$,

называется ***вероятностным пространством***, *а мера* $\mathbf{P}$ — ***вероятностной мерой*** *или просто* ***вероятностью***.

Замечания: 1. Предполагается, что $\{\Omega, \mathcal{F}, \mathbf{P}\}$ — полное вероятностное пространство (см. разд. 4.1.2).

2. Над случайными событиями из $\mathcal{F}$ можно совершать действия, аналогичные действиям над множествами, для которых мы будем использовать следующие обозначения:

$$A + B = A \cup B, \quad AB = A \cap B, \quad A \setminus B, \quad \overline{A} = \Omega \setminus A,$$

$$\sum_k A_k = \bigcup_k A_k, \quad \prod_k A_k = \bigcap_k A_k,$$

где событие $\overline{A}$ называется ***противоположным*** A.

Вероятность $\mathbf{P}\{\cdot\}$ обладает следующими свойствами:

1) $0 \le \mathbf{P}\{A\} \le 1$ для любого события $A \in \mathcal{F}$;

2) $\mathbf{P}\{\Omega\} = 1$, где Ω — ***достоверное событие***;

3) $\mathbf{P}\{\varnothing\} = 0$, где $\varnothing = \overline{\Omega}$ — ***невозможное событие***;

4) $\mathbf{P}\{A+B\} = \mathbf{P}\{A\} + \mathbf{P}\{B\}$, если $AB = \varnothing$, т. е. события A и B ***несовместны***;

5) $\mathbf{P}\{A+B\} = \mathbf{P}\{A\} + \mathbf{P}\{B\} - \mathbf{P}\{AB\}$ для любых событий $A, B \in \mathcal{F}$;

6) $\mathbf{P}\{A\} \le \mathbf{P}\{B\}$, если $A \subseteq B$ (т. е. A — ***частный случай*** события B).

Замечание. Указанные свойства следуют из общих свойств меры (см. разд. 4.1.2). Свойство 4 распространяется очевидным образом на любое конечное или счетное множество несовместных событий:

$$\mathbf{P}\left\{\sum_{k=1}^{\infty} A_k\right\} = \sum_{k=1}^{\infty} \mathbf{P}\{A_k\}, \quad A_k \in \mathcal{F}, \quad A_m A_n = \varnothing \text{ при } m \ne n.$$

Определение 4.2.2. *События A и B называются* ***независимыми****, если $\mathbf{P}\{AB\} = \mathbf{P}\{A\}\,\mathbf{P}\{B\}$. События $\{A_n\}$* ***независимы в совокупности****, если для любого конечного набора событий A_{n_k}, $k = 1, \dots, m$*

$$\mathbf{P}\left\{\prod_{k=1}^{m} A_{n_k}\right\} = \prod_{k=1}^{m} \mathbf{P}\{A_{n_k}\},$$

где m может равняться ∞.

Определение 4.2.3. ***Условной вероятностью*** *события A относительно события B такого, что $\mathbf{P}\{B\} > 0$, называется величина*

$$\mathbf{P}\{A \mid B\} = \frac{\mathbf{P}\{AB\}}{\mathbf{P}\{B\}}.$$

Если события A, B независимы и имеют положительные вероятности, то $\mathbf{P}\{A \mid B\} = \mathbf{P}\{A\}$ и $\mathbf{P}\{B \mid A\} = \mathbf{P}\{B\}$.

Пусть события $H_1, \ldots, H_N \in \mathcal{F}$ удовлетворяют условиям:

а) $\mathbf{P}\{H_k\} > 0$ при всех k;

б) $H_m H_n = \varnothing$, если $m \neq n$;

в) $\sum\limits_{k=1}^{N} H_k = \Omega$,

тогда для любого $A \in \mathcal{F}$ справедлива ***формула полной вероятности***:

$$\mathbf{P}\{A\} = \sum_{k=1}^{N} \mathbf{P}\{H_k\}\mathbf{P}\{A \mid H_k\}.$$

События $\{H_k\}$ обычно называют ***вероятностными гипотезами***.

4.2.2. Случайные величины и векторы

Определение 4.2.4. ***Случайной величиной (СВ)****, определенной на $\{\Omega, \mathcal{F}, \mathbf{P}\}$, называется числовая функция $\xi(\omega)$, $\omega \in \Omega$, измеримая относительно $\mathcal{F}$.*

Определение 4.2.4 означает, что для всякого борелевского подмножества $B \subseteq \mathbb{R}^1$ множество

$$\xi^{-1}(B) = \{\omega \in \Omega : \xi(\omega) \in B\} \tag{4.2.1}$$

является случайным событием.

Далее для краткости будем опускать аргумент ω: ξ, $\{\xi \in B\}$ и т. п.

В силу теорем разд. 4.1.4 сумма, разность, произведение и частное двух случайных величин (при условии, что знаменатель не обращается в нуль) также являются случайными величинами.

Определение 4.2.5. *Две случайные величины ξ и η, заданные на $\{\Omega, \mathcal{F}, \mathbf{P}\}$, называются* ***эквивалентными****, если*

$$\mathbf{P}\{\xi \neq \eta\} = 0.$$

Если некоторое утверждение относительно СВ ξ (или совокупности СВ) выполнено для всех $\omega \in \Omega \setminus A$, причем $\mathbf{P}\{A\} = 0$, то говорят, что это утверждение выполнено ***почти наверное*** по мере $\mathbf{P}$ (или ***с вероятностью 1***). Мы будем сопровождать соответствующее утверждение знаком ($\mathbf{P}$-п.н.). Например, если ξ и η эквивалентны, то $\xi = \eta$ ($\mathbf{P}$-п.н.).

Из теоремы 4.1.4 следует, что если ξ — случайная величина, а $g(x)$, $x \in \mathbb{R}^1$ — борелевская функция, то $g(\xi)$ также является случайной величиной, так как функция $\eta(\omega) = g(\xi(\omega))$, $\omega \in \Omega$ является $\mathcal{F}$-измеримой. В частности, любое непрерывное преобразование СВ ξ приводит к СВ η (т. е. свойство $\mathcal{F}$-измеримости сохраняется).

Определение 4.2.6. *Пусть ξ — некоторая случайная величина. Наименьшую σ-алгебру, содержащую события вида*

$$\xi^{-1}(B) = \{\xi \in B\}, \quad B \in \mathcal{B}(\mathbb{R}^1), \tag{4.2.2}$$

будем называть ***σ-алгеброй, порожденной случайной величиной*** *ξ. Эта σ-алгебра обозначается $\mathcal{F}^\xi$, а также $\sigma\{\xi\}$.*

По определению $\mathcal{F}^\xi \subseteq \mathcal{F}$. Введенная σ-алгебра $\mathcal{F}^\xi$ состоит из всех случайных событий $A \in \mathcal{F}$, о наступлении которых мы можем судить, наблюдая СВ ξ.

Теорема 4.2.1. *Пусть $\varphi(x)$ — произвольная борелевская функция, тогда СВ $\eta = \varphi(\xi)$ является $\mathcal{F}^\xi$-измеримой. Наоборот, если некоторая СВ η является $\mathcal{F}^\xi$-измеримой, то существует борелевская функция $\varphi(x)$ такая, что $\eta = \varphi(\xi)$.*

Определение 4.2.7. *Функция $F_\xi(x) = \mathbf{P}\{\xi \le x\}$, $x \in \mathbb{R}^1$ называется* ***функцией распределения*** *СВ ξ.*

Функция распределения $F_\xi(x)$ обладает всеми свойствами функции распределения меры на измеримом пространстве $\{\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)\}$ (см. определение 4.1.9, которое полностью применимо с учетом соотношения $\mu(\mathbb{R}^1) = 1$).

Для любой функции распределения $F(x)$ существует полное вероятностное пространство $\{\Omega, \mathcal{F}, \mathbf{P}\}$ и заданная на нем СВ ξ такая, что $F_\xi(x) = F(x)$.

Рассмотрим конкретные типы функций распределения.

Определение 4.2.8. *Если СВ ξ принимает значения из конечного или счетного множества $\{a_1, \dots, a_n, \dots\}$ с вероятностями соответственно $\{p_1, \dots, p_n, \dots\}$, где $p_n > 0$, $\sum\limits_n p_n = 1$, то говорят, что случайная величина является* ***дискретной*** *(или имеет* ***дискретное распределение****). Ее* *функция распределения имеет вид*

$$F_\xi(x) = \sum_{k:\, a_k \le x} p_k, \quad x \in \mathbb{R}^1.$$

Таким образом, функция распределения дискретной СВ имеет разрывы первого рода в точках a_k, а величины скачков равны $F_\xi(a_k) - F_\xi(a_k-) = p_k$. При этом для любого множества $B \in \mathcal{B}(\mathbb{R}^1)$

$$\mathbf{P}\{\xi \in B\} = \sum_{k:\, a_k \in B} p_k.$$

Если множество значений, которые принимает дискретная СВ, конечно, то ее функция распределения кусочно-постоянна.

Функция множества $\mathbf{P}_\xi(B) = \mathbf{P}\{\xi \in B\}$ является ***мерой*** на $\mathcal{B}(\mathbb{R}^1)$ и называется ***законом распределения СВ*** ξ. При этом $F_\xi(x)$ является функцией распределения этой меры (см. разд. 4.1.3). При определенных условиях мера $\mathbf{P}_\xi(B)$ абсолютно непрерывна относительно меры Лебега на $\mathcal{B}(\mathbb{R}^1)$, т. е. $F_\xi(x)$ имеет плотность распределения $p_\xi(x)$.

Определение 4.2.9. *Если функция распределения $F_\xi(x)$ случайной величины ξ допускает представление*

$$F_\xi(x) = \int_{-\infty}^{x} p_\xi(y)dy, \quad \text{где} \quad p_\xi(y) \ge 0, \quad \int_{-\infty}^{\infty} p_\xi(y)dy = 1,$$

*а интеграл понимается как интеграл Лебега, то СВ ξ называется **абсолютно непрерывной** (имеет **непрерывное распределение**). Для любого множества $B \in \mathcal{B}(\mathbb{R}^1)$*

$$\mathbf{P}_\xi(B) = \int_B p_\xi(y)dy.$$

*Функция $p_\xi(y)$ называется **плотностью распределения** СВ ξ.*

Заметим, что функция $F_\xi(x)$ в данном случае не имеет разрывов и почти всюду на $\mathbb{R}^1$ дифференцируема: $\dfrac{dF_\xi(x)}{dx} = p_\xi(x)$.

В общем случае функция распределения $F_\xi(x)$ непрерывна справа в каждой точке разрыва $x \in \mathbb{R}^1$. Для вычисления вероятности события $\{\xi \in B\}$, где $B \in \mathcal{B}(\mathbb{R}^1)$, следует вычислять интеграл Лебега–Стилтьеса:

$$\mathbf{P}\{\xi \in B\} = \int_B dF_\xi(y) = \mathbf{P}_\xi(B), \quad B \in \mathcal{B}(\mathbb{R}^1).$$

Предположим, что СВ $\xi_1, \ldots, \xi_n$ определены на одном вероятностном пространстве $\{\Omega, \mathcal{F}, \mathbf{P}\}$. Тогда упорядоченный набор n случайных величин $\xi = \{\xi_1, \ldots, \xi_n\}^*$ будем называть ***n*-мерным случайным вектором**.

Определение 4.2.10. *Наименьшую σ-алгебру, содержащую все события вида*

$$\xi^{-1}(B) = \{\xi \in B\}, \quad B \in \mathcal{B}(\mathbb{R}^n), \tag{4.2.3}$$

*будем называть **σ-алгеброй, порожденной случайным вектором** ξ. Эта σ-алгебра обозначается $\sigma\{\xi\}$ или $\mathcal{F}^{\xi}$.*

Определение 4.2.11. *Функция*

$$F_{\xi}(x_1, \ldots, x_n) = \mathbf{P}\{(\xi_1 \le x_1) \cdot \ldots \cdot (\xi_n \le x_n)\}, \quad x_1, \ldots, x_n \in \mathbb{R}^1$$

*называется **функцией распределения *n*-мерного случайного вектора** ξ.*

Функция распределения случайного вектора обладает всеми свойствами функции распределения меры на измеримом пространстве $\{\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\}$ (см. определение 4.1.12), которое полностью применимо с учетом соотношения $\mu(\mathbb{R}^n) = 1$.

Для любого борелевского множества $B \in \mathcal{B}(\mathbb{R}^n)$ вероятность случайного события $\{\xi \in B\}$ вычисляется как интеграл Лебега:

$$\mathbf{P}_{\xi}(B) = \mathbf{P}\{\xi \in B\} = \int\limits_B dF_{\xi}(x_1, \ldots, x_n).$$

Мера $\mathbf{P}_{\xi}(\cdot)$ на $\mathcal{B}(\mathbb{R}^n)$ — ***закон распределения случайного вектора*** ξ, заданный с помощью $F_{\xi}(x_1, \ldots, x_n)$. Если мера $\mathbf{P}_{\xi}(\cdot)$ имеет плотность, т. е.

$$\mathbf{P}\{\xi \in B\} = \int\limits_B p_{\xi}(y_1, \ldots, y_n) dy_1 \ldots dy_n \quad \forall B \in \mathcal{B}(\mathbb{R}^n),$$

то

$$F_{\xi}(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} p_{\xi}(y_1, \ldots, y_n) dy_1 \ldots dy_n.$$

Определение 4.2.12. *Случайные величины $\{\xi_1, \ldots, \xi_n\}$ **независимы в совокупности**, если для любого набора множеств $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}^1)$ события $\{\xi_1 \in B_1\}, \ldots, \{\xi_n \in B_n\}$ независимы в совокупности.*

Теорема 4.2.2. *Для того чтобы случайные величины $\{\xi_1, \ldots, \xi_n\}$ были независимы в совокупности, необходимо и достаточно, чтобы для всех $x_1, \ldots, x_n \in \mathbb{R}^1$*

$$F_\xi(x_1, \ldots, x_n) = \prod_{k=1}^{n} F_{\xi_k}(x_k).$$

Общий способ определения независимости случайных событий и величин состоит в следующем: пусть $\mathcal{F}_1 \subseteq \mathcal{F}$ и $\mathcal{F}_2 \subseteq \mathcal{F}$ — некоторые σ-алгебры случайных событий. ***Системы $\mathcal{F}_1$ и $\mathcal{F}_2$ независимы***, если независимы любые два события $A \in \mathcal{F}_1$ и $B \in \mathcal{F}_2$, т. е. $\mathbf{P}\{AB\} = \mathbf{P}\{A\}\mathbf{P}\{B\}$.

Теперь нетрудно определить понятие независимости СВ ξ и η: ξ и η независимы тогда и только тогда, когда независимы σ-алгебры $\mathcal{F}^\xi$ и $\mathcal{F}^\eta$. В частности, СВ ξ не зависит от случайного события A, если A и $\mathcal{F}^\xi$ независимы (т. е. независимы A и любое $B \in \mathcal{F}^\xi$). Эти понятия будут использованы в дальнейшем при построении условного математического ожидания.

В общем случае случайная величина ξ может принимать ***комплексные значения***, т. е. $\xi = \alpha + i\beta$, где α, β — СВ, заданные на $\{\Omega, \mathcal{F}, \mathbf{P}\}$, а $i^2 = -1$. Изучение комплексной СВ фактически сводится к изучению двумерного случайного вектора $\{\alpha, \beta\}^*$.

Приведем примеры некоторых наиболее важных законов распределения.

1. Дискретная СВ ξ имеет ***биномиальное распределение*** с параметрами $(N; p)$, где $0 < p < 1$, и обозначается $\mathrm{Bi}(N; p)$, если

$$\mathbf{P}\{\xi = m\} = C_N^m p^m q^{N-m}, \quad m = 0, 1, \ldots, N,$$

где $C_N^m = \dfrac{N!}{m!(N-m)!}$ — число сочетаний из N по m, $q = 1-p$. Распределение $\mathrm{Bi}(1; p)$ называется ***распределением Бернулли***.

2. Дискретная СВ ξ имеет ***распределение Пуассона*** с параметром $\lambda > 0$, и обозначается $\Pi(\lambda)$, если

$$\mathbf{P}\{\xi = m\} = \frac{\lambda^m}{m!}\, e^{-\lambda}, \quad m = 0, 1, \ldots$$

3. Непрерывная СВ ξ имеет ***равномерное распределение*** на отрезке $[a, b]$, и обозначается $\mathcal{R}(a, b)$, если ее плотность распределения имеет вид

$$p_\xi(x) = \begin{cases} 1/(b-a), & \text{если } x \in [a, b], \\ 0, & \text{если } x \notin [a, b]. \end{cases}$$

4. Непрерывная СВ ξ имеет ***экспоненциальное*** (или ***показательное***) ***распределение*** с параметром $\lambda > 0$, и обозначается $E(\lambda)$, если

$$p_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

5. Непрерывная СВ ξ имеет ***гауссовское*** (или ***нормальное***) ***распределение*** с параметрами $(m; \sigma^2)$, где $\sigma > 0$, если

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}^1.$$

Для гауссовского распределения используется обозначение $\mathcal{N}(m; \sigma^2)$.

Числовые характеристики этих распределений приведены в следующем разделе.

4.2.3. Математическое ожидание

Определение 4.2.13. ***Математическим ожиданием*** *(или* ***средним****) СВ ξ, определенной на $\{\Omega, \mathcal{F}, \mathbf{P}\}$, называется число*

$$\mathbf{M}\{\xi\} = m_\xi = \int_\Omega \xi(\omega) \mathbf{P}\{d\omega\}\,. \tag{4.2.4}$$

Математическое ожидание определено, если интеграл Лебега в правой части равенства (4.2.4) существует.

Таким образом, математическое ожидание СВ ξ есть интеграл Лебега от функции $\xi(\omega)$ на Ω по вероятностной мере $\mathbf{P}$. Если $\mathbf{P}_\xi(\cdot)$ — закон распределения СВ ξ на $\mathcal{B}(\mathbb{R}^1)$, а $F_\xi(y)$ — соответствующая функция распределения, то $\mathbf{M}\{\xi\}$ можно вычислить следующим образом:

$$\mathbf{M}\{\xi\} = \int_{\mathbb{R}^1} y\, \mathbf{P}_\xi(dy) = \int_{-\infty}^{\infty} y\, dF_\xi(y), \tag{4.2.5}$$

причем первый интеграл понимается как интеграл Лебега по мере $\mathbf{P}_\xi(\cdot)$, а второй — как интеграл Лебега–Стилтьеса. Существование интегралов в правой части (4.2.5) вытекает из существования математического ожидания, и, наоборот, из существования интегралов следует существование математического ожидания. Формула (4.2.5) следует из (4.2.4) и теоремы 4.1.15 о замене переменной в интеграле Лебега.

Если функция распределения $F_\xi(x)$ является комбинацией абсолютно-непрерывной и дискретной составляющих, т. е. допускает представление вида $F_\xi(x) = \int\limits_{-\infty}^{x} p_\xi(y)dy + \sum\limits_{k:\, a_k \le x} p_k$, где $p_\xi(y) \ge 0$, а $p_k > 0$ — величина скачка в точке разрыва a_k, то (4.2.5) принимает вид

$$\mathbf{M}\{\xi\} = \int\limits_{-\infty}^{\infty} y p_\xi(y)dy + \sum_k p_k a_k.$$

СВ ξ называется ***центрированной***, если $\mathbf{M}\{\xi\} = 0$.

Основные свойства математического ожидания вытекают из свойств интеграла Лебега (см. разд. 4.1.5).

1. Математические ожидания $\mathbf{M}\{\xi\}$ и $\mathbf{M}\{|\xi|\}$ существуют или не существуют одновременно, причем $|\mathbf{M}\{\xi\}\,| \le \mathbf{M}\{|\xi|\}$.

2. $\mathbf{M}\{I_A\} = \int\limits_{\Omega} I_A(\omega)\mathbf{P}\{d\omega\} = \mathbf{P}\{A\}$, где I_A — индикатор события $A \in \mathcal{F}$.

3. Если $\mathbf{M}\{\xi\}$ существует, то для любой константы λ

$$\mathbf{M}\{\lambda\xi\} = \lambda\mathbf{M}\{\xi\}\,.$$

4. Если $\mathbf{M}\{\xi\}$ и $\mathbf{M}\{\eta\}$ существуют, то

$$\mathbf{M}\{\xi + \eta\} = \mathbf{M}\{\xi\} + \mathbf{M}\{\eta\}\,.$$

5. Если $\xi \le \eta$, то $\mathbf{M}\{\xi\} \le \mathbf{M}\{\eta\}$.

6. Если $\varphi(x)$ — борелевская функция, то

$$\mathbf{M}\{\varphi(\xi)\} = \int\limits_{-\infty}^{\infty} \varphi(x)dF_\xi(x),$$

где математическое ожидание и интеграл существуют или не существуют одновременно.

7. Если $\mathbf{M}\{\xi\}$ определено, то для каждого $A \in \mathcal{F}$ существует

$$\mathbf{M}\{\xi I_A\} = \int\limits_A \xi(\omega)\mathbf{P}(d\omega).$$

8. Если $\xi \geq 0$ ($\mathbf{P}$-п.н.), а также $\mathbf{M}\{\xi\} < \infty$, то функция множества $Q\{A\} = \mathbf{M}\{\xi I_A\}$, $A \in \mathcal{F}$ является конечной мерой на $\mathcal{F}$.

9. Если $\xi \geq 0$ ($\mathbf{P}$-п.н.), $\mathbf{M}\{\xi\} < \infty$ и $\varepsilon > 0$, то

$$\mathbf{P}\{\xi \geq \varepsilon\} \leq \frac{\mathbf{M}\{\xi\}}{\varepsilon}.$$

10. Если $\mathbf{M}\{|\xi|^p\} < \infty$, $p > 0$, то выполняется ***неравенство Маркова***

$$\mathbf{P}\{|\xi| \geq \varepsilon\} \leq \frac{\mathbf{M}\{|\xi|^p\}}{\varepsilon^p}.$$

Определение 4.2.14. ***Дисперсией*** $\mathbf{D}\{\xi\}$ *СВ* ξ *называется число*

$$\mathbf{D}\{\xi\} = D_\xi = \mathbf{M}\big\{|\xi - m_\xi|^2\big\} = \int\limits_{-\infty}^{\infty} |y - m_\xi|^2 dF_\xi(y).$$

Из определения и свойств интеграла Лебега следует:

1) $\mathbf{D}\{\xi\} \geq 0$;

2) $\mathbf{D}\{a\xi + b\} = |a|^2 \mathbf{D}\{\xi\}$, если $a, b = \mathrm{const}$;

3) $\mathbf{D}\left\{\sum\limits_{k=1}^{n} \xi_k\right\} = \sum\limits_{k=1}^{n} \mathbf{D}\{\xi_k\}$, если $\mathbf{D}\{\xi_k\} < \infty$, а СВ $\{\xi_k\}$ — независимы в совокупности;

4) $\mathbf{D}\{\xi\} = \mathbf{M}\big\{|\xi|^2\big\} - |m_\xi|^2$;

5) если $\mathbf{M}\big\{|\xi|^2\big\} < \infty$, то выполнено ***неравенство Чебышева***

$$\mathbf{P}\{|\xi - m_\xi| \geq \varepsilon\} \leq \frac{\mathbf{D}\{\xi\}}{\varepsilon^2}.$$

Определение 4.2.15. ***Ковариацией*** *СВ* ξ *и* η *называется величина*

$$\mathbf{cov}\{\xi, \eta\} = \mathbf{M}\Big\{(\xi - m_\xi)\overline{(\eta - m_\eta)}\Big\},$$

где $\overline{(\cdot)}$ *— знак комплексного сопряжения.*

Перечислим важнейшие свойства ковариации:

1) если $\mathbf{M}\big\{|\xi|^2\big\} < \infty$, $\mathbf{M}\big\{|\eta|^2\big\} < \infty$, то ковариация СВ ξ, η существует и удовлетворяет ***неравенству Коши–Буняковского***:

$$|\mathbf{cov}\{\xi, \eta\}|^2 \leq \mathbf{D}\{\xi\}\,\mathbf{D}\{\eta\};$$

2) если $\mathbf{M}\{\xi\} = \mathbf{M}\{\eta\} = 0$, то $\mathbf{cov}\{\xi, \eta\} = (\xi, \eta)$ — скалярное произведение случайных величин ξ, η;

3) если ξ и η независимы, то $\mathbf{cov}\{\xi, \eta\} = 0$;

4) $\mathbf{D}\{\xi\} = \mathbf{cov}\{\xi, \xi\}$;

5) $\mathbf{D}\{\xi + \eta\} = \mathbf{D}\{\xi\} + \mathbf{D}\{\eta\} + 2\,\mathbf{cov}\{\xi, \eta\}$.

Если $\mathbf{cov}\{\xi, \eta\} = 0$, то СВ ξ и η называются ***некоррелированными*** или ***ортогональными***, что обозначается $\xi \perp \eta$.

Приведем числовые характеристики (т. е. математические ожидания и дисперсии) случайных величин, рассмотренных в конце разд. 4.2.2:

1) биномиальное распределение с параметрами $(N; p)$:

$$m_\xi = Np, \quad D_\xi = Npq;$$

2) распределение Пуассона с параметром $\lambda > 0$:

$$m_\xi = D_\xi = \lambda;$$

3) равномерное распределение на отрезке $[a, b]$:

$$m_\xi = \frac{a+b}{2}, \quad D_\xi = \frac{(b-a)^2}{12};$$

4) экспоненциальное распределение с параметром $\lambda > 0$:

$$m_\xi = \frac{1}{\lambda}, \quad D_\xi = \frac{1}{\lambda^2};$$

5) гауссовское распределение с параметрами $(m; \sigma^2)$, $\sigma > 0$:

$$m_\xi = m, \quad D_\xi = \sigma^2.$$

4.2.4. Последовательности случайных величин

Будем далее предполагать, что случайные величины $\{\xi_n,\ n = 1, 2, \ldots\}$ заданы на одном и том же вероятностном пространстве $\{\Omega, \mathcal{F}, \mathbf{P}\}$.

Определение 4.2.16. *Последовательность случайных величин* $\{\xi_n\}$ *называется* ***сходящейся почти наверное*** ($\mathbf{P}$-п.н.) *к СВ* ξ, *если*

$$\mathbf{P}\{\omega \in \Omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\} = 1.$$

В этом случае используются обозначения: $\xi_n \to \xi$ ($\mathbf{P}$-п.н.) или $\xi_n \xrightarrow{\text{п.н.}} \xi$, $n \to \infty$, а также говорят, что ξ_n сходится к ξ ***с вероятностью 1***.

Следующие результаты вытекают из теорем о предельном переходе под знаком интеграла Лебега и задают правила предельного перехода под знаком математического ожидания.

Теорема 4.2.3. *Если последовательность $\{\xi_n\}$ сходится к ξ ($\mathbf{P}$-п.н.) и найдется СВ η такая, что $|\xi_n| \le \eta$ ($\mathbf{P}$-п.н.) для всякого n и $\mathbf{M}\{\eta\} < \infty$, то*

$$\mathbf{M}\{|\xi|\} < \infty, \quad \mathbf{M}\{\xi_n\} \to \mathbf{M}\{\xi\}$$

и

$$\mathbf{M}\{|\xi_n - \xi|\} \to 0 \quad \text{при} \quad n \to \infty.$$

Теорема 4.2.4. *Пусть $\eta, \xi_1, \xi_2, \ldots$ — случайные величины.*

1. *Если $\xi_n \ge \eta$ для всех $n \ge 1$, где $\mathbf{M}\{\eta\} > -\infty$, и $\xi_n \uparrow \xi$ ($\mathbf{P}$-п.н.), то $\mathbf{M}\{\xi_n\} \uparrow \mathbf{M}\{\xi\}$ при $n \to \infty$.*

2. *Если $\xi_n \le \eta$ для всех $n \ge 1$, где $\mathbf{M}\{\eta\} < \infty$, и $\xi_n \downarrow \xi$ ($\mathbf{P}$-п.н.), то $\mathbf{M}\{\xi_n\} \downarrow \mathbf{M}\{\xi\}$ при $n \to \infty$.*

Теорема 4.2.5. *Если последовательность $\{\xi_n\}$ неотрицательных случайных величин сходится ($\mathbf{P}$-п.н.) к СВ ξ, причем существует такая константа $K < \infty$, что $\mathbf{M}\{\xi_n\} \le K$ для всех n. Тогда $\mathbf{M}\{\xi\}$ существует и $\mathbf{M}\{\xi\} \le K$.*

Определение 4.2.17. *Последовательность случайных величин $\{\xi_n\}$ называется* ***сходящейся по вероятности*** *к СВ ξ, если для любого $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathbf{P}\{|\xi_n - \xi| > \varepsilon\} = 0.$$

Теорема 4.2.6. *У любой последовательности случайных величин $\{\xi_n, n = 1, 2, \ldots\}$, сходящейся по вероятности к ξ, найдется подпоследовательность $\{\xi_{n_k}\}$ такая, что $\xi_{n_k} \xrightarrow{\text{п.н.}} \xi$ при $k \to \infty$.*

Определение 4.2.18. *Последовательность случайных величин $\{\xi_n\}$ называется* ***фундаментальной по вероятности****, если для любого $\varepsilon > 0$*

$$\lim_{n,m\to\infty} \mathbf{P}\{|\xi_n - \xi_m| > \varepsilon\} = 0.$$

Теорема 4.2.7. *Для сходимости последовательности $\{\xi_n\}$ по вероятности необходимо и достаточно, чтобы последовательность была фундаментальной.*

Аналогичный критерий существует и для сходимости ($\mathbf{P}$-п.н.).

Теорема 4.2.8. *Для сходимости последовательности* $\{\xi_n\}$ ($\mathbf{P}$-п.н.) *необходимо и достаточно, чтобы последовательность была* ***фундаментальной*** ($\mathbf{P}$-п.н.), *т. е. для любого* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}\Big\{\sup_{k\geq 0} |\xi_{n+k} - \xi_n| > \varepsilon\Big\} = 0.$$

Определение 4.2.19. *Последовательность случайных величин* $\{\xi_n\}$ *называется* ***сходящейся в среднем порядка*** $p > 0$ *к СВ* ξ, *если*

$$\lim_{n\to\infty} \mathbf{M}\{|\xi_n - \xi|^p\} = 0.$$

При $p = 2$ этот вид сходимости называется сходимостью ***в среднем квадратическом*** (или ***с.к.-сходимостью***) и обозначается $\xi = \underset{n\to\infty}{\mathrm{l.i.m.}}\,\xi_n$ или $\xi_n \xrightarrow{\text{с.к.}} \xi$, $n \to \infty$. Среднеквадратическая сходимость имеет особое значение в теории случайных процессов.

Приведем важнейшие свойства с.к.-сходимости.

1. ***Критерий Коши.*** Для того чтобы существовал с.к.-предел ξ последовательности СВ $\{\xi_n\}$ при $n \to \infty$, необходимо и достаточно, чтобы $\{\xi_n\}$ была фундаментальной в с.к.-смысле:

$$\mathbf{M}\big\{|\xi_n - \xi_m|^2\big\} \to 0 \quad \text{при} \quad n, m \to \infty.$$

2. Пусть $\xi_n \xrightarrow{\text{с.к.}} \xi$, $n \to \infty$, причем $\mathbf{M}\big\{|\xi_n|^2\big\} < \infty$, тогда $\mathbf{M}\big\{|\xi|^2\big\} < \infty$ и

$$\lim_{n\to\infty} \mathbf{M}\{\xi_n\} = \mathbf{M}\Big\{\underset{n\to\infty}{\mathrm{l.i.m.}}\,\xi_n\Big\} = \mathbf{M}\{\xi\}\,;$$
$$\lim_{n\to\infty} \mathbf{M}\big\{|\xi_n|^2\big\} = \mathbf{M}\Big\{|\underset{n\to\infty}{\mathrm{l.i.m.}}\,\xi_n|^2\Big\} = \mathbf{M}\big\{|\xi|^2\big\}\,.$$

3. Если последовательности СВ $\{\xi_n\}$, $\{\eta_n\}$ таковы, что $\mathbf{M}\big\{|\xi_n|^2\big\} < \infty$, $\mathbf{M}\big\{|\eta_n|^2\big\} < \infty$ и $\xi_n \xrightarrow{\text{с.к.}} \xi$, $\eta_n \xrightarrow{\text{с.к.}} \eta$ при $n \to \infty$, то

$$\mathbf{M}\{\xi_n\overline{\eta}_n\} \to \mathbf{M}\{\xi\overline{\eta}\} \quad \text{при} \quad n \to \infty.$$

4. ***Лемма Лоэва.*** Последовательность СВ $\{\xi_n\}$ имеет с.к.-предел тогда и только тогда, когда существует неотрицательное число c такое, что $\mathbf{M}\big\{\xi_n\overline{\xi}_m\big\} \to c$ при $n, m \to \infty$. При этом $c = \mathbf{M}\big\{|\xi|^2\big\}$, где $\xi = \underset{n\to\infty}{\mathrm{l.i.m.}}\,\xi_n$.

Наиболее слабый вид сходимости — сходимость последовательности по распределению.

Определение 4.2.20. *Последовательность* $\{\xi_n,\ n = 1, 2, \ldots\}$ *случайных величин называется* ***сходящейся по распределению*** *(или* ***слабо сходящийся****) к СВ* ξ, *если для любой равномерно ограниченной непрерывной функции* $f(x)$, $x \in \mathbb{R}^1$

$$\lim_{n\to\infty} \mathbf{M}\{f(\xi_n)\} = \mathbf{M}\{f(\xi)\}\,.$$

Соотношения между всеми типами сходимости описываются теоремой.

Теорема 4.2.9. *Справедливы следующие утверждения:*

1. *"Сходимость* ($\mathbf{P}$-п.н.)*"* $\Longrightarrow$ *"Сходимость по вероятности".*

2. *"Сходимость в среднем порядка* $p > 0$*"* $\Longrightarrow$ *"Сходимость по вероятности".*

3. *"Сходимость по вероятности"* $\Longrightarrow$ *"Сходимость по распределению".*

4. *Если* ξ_n *слабо сходится к константе* a, *то* ξ_n *сходится к* a *и по вероятности.*

4.2.5. Условное математическое ожидание

Пусть $\{\Omega, \mathcal{F}, \mathbf{P}\}$ — заданное вероятностное пространство, $\mathcal{G}$ — некоторая σ-алгебра случайных событий, т. е. $\mathcal{G} \subseteq \mathcal{F}$, и ξ — СВ такая, что $\mathbf{M}\{|\xi|\} < \infty$.

Определение 4.2.21. ***Условным математическим ожиданием СВ*** ξ ***относительно*** $\mathcal{G}$ *называется случайная величина, обозначаемая* $\mathbf{M}\{\xi \mid \mathcal{G}\}$ *и удовлетворяющая следующим условиям:*

1) $\mathbf{M}\{\xi \mid \mathcal{G}\}$ *является* $\mathcal{G}$-*измеримой;*

2) *для любого множества* $A \in \mathcal{G}$ *выполняется равенство*

$$\int_A \xi(\omega)\,\mathbf{P}\{d\omega\} = \int_A \mathbf{M}\{\xi \mid \mathcal{G}\}\,(\omega)\,\mathbf{P}\{d\omega\}\,. \tag{4.2.6}$$

Теорема 4.2.10. *Условное математическое ожидание СВ* ξ *определено единственным образом* ($\mathbf{P}$-п.н.).

Свойства условного математического ожидания.

Следующие свойства условного математического ожидания непосредственно вытекают из определения 4.2.21.

1. Если $\xi = C = \text{const}$ ($\mathbf{P}$-п.н.), то $\mathbf{M}\{\xi \mid \mathcal{G}\} = C$ ($\mathbf{P}$-п.н.).

2. Если $\xi \le \eta$ ($\mathbf{P}$-п.н.), то $\mathbf{M}\{\xi \mid \mathcal{G}\} \le \mathbf{M}\{\eta \mid \mathcal{G}\}$ ($\mathbf{P}$-п.н.).

3. $|\mathbf{M}\{\xi \mid \mathcal{G}\}| \le \mathbf{M}\{|\xi| \mid \mathcal{G}\}$ ($\mathbf{P}$-п.н.).

4. Если a, b — константы, а ξ, η — случайные величины такие, что $\mathbf{M}\{|\xi|\} < \infty$, $\mathbf{M}\{|\eta|\} < \infty$, то

$$\mathbf{M}\{a\xi + b\eta \mid \mathcal{G}\} = a\,\mathbf{M}\{\xi \mid \mathcal{G}\} + b\,\mathbf{M}\{\eta \mid \mathcal{G}\} \quad (\mathbf{P}\text{-п.н.}).$$

5. Пусть $\mathcal{G} = \{\varnothing, \Omega\}$ — тривиальная σ-алгебра, тогда

$$\mathbf{M}\{\xi \mid \mathcal{G}\} = \mathbf{M}\{\xi\} \quad (\mathbf{P}\text{-п.н.}).$$

6. Пусть СВ ξ измерима относительно $\mathcal{G}$, тогда $\mathbf{M}\{\xi \mid \mathcal{G}\} = \xi$ ($\mathbf{P}$-п.н.).

7. $\mathbf{M}\{\mathbf{M}\{\xi \mid \mathcal{G}\}\} = \mathbf{M}\{\xi\}$ ($\mathbf{P}$-п.н.).

8. Если $\mathcal{G}_1 \subseteq \mathcal{G}_2$, то

$$\mathbf{M}\{\mathbf{M}\{\xi \mid \mathcal{G}_2\} \mid \mathcal{G}_1\} = \mathbf{M}\{\xi \mid \mathcal{G}_1\} \quad (\mathbf{P}\text{-п.н.}).$$

9. Если $\mathcal{G}_2 \subseteq \mathcal{G}_1$, то

$$\mathbf{M}\{\mathbf{M}\{\xi \mid \mathcal{G}_2\} \mid \mathcal{G}_1\} = \mathbf{M}\{\xi \mid \mathcal{G}_2\} \quad (\mathbf{P}\text{-п.н.}).$$

10. Если СВ ξ не зависит от $\mathcal{G}$ (т. е. $\mathcal{F}^{\xi}$ не зависит от $\mathcal{G}$), то

$$\mathbf{M}\{\xi \mid \mathcal{G}\} = \mathbf{M}\{\xi\} \quad (\mathbf{P}\text{-п.н.}).$$

11. Пусть η измерима относительно $\mathcal{G}$ и $\mathbf{M}\{|\xi\eta|\} < \infty$, тогда

$$\mathbf{M}\{\xi\eta \mid \mathcal{G}\} = \eta\,\mathbf{M}\{\xi \mid \mathcal{G}\} \quad (\mathbf{P}\text{-п.н.}).$$

12. ***Неравенство Иенсена.*** Пусть $g(x)$ — выпуклая вниз функция такая, что $\mathbf{M}\{|g(x)|\} < \infty$, тогда

$$g\left(\mathbf{M}\{\xi \mid \mathcal{G}\}\right) \le \mathbf{M}\{g(\xi) \mid \mathcal{G}\} \quad (\mathbf{P}\text{-п.н.}).$$

13. Пусть η — произвольная $\mathcal{G}$-измеримая случайная величина. Если $\mathbf{M}\big\{|\xi|^2\big\} < \infty$, $\mathbf{M}\big\{|\eta|^2\big\} < \infty$, то

$$\mathbf{M}\big\{|\xi - \mathbf{M}\{\xi \mid \mathcal{G}\}\,|^2\big\} \le \mathbf{M}\big\{|\xi - \eta|^2\big\} \quad (\mathbf{P}\text{-п.н.}).$$

Определение 4.2.22. *Случайная величина* $\mathbf{M}\{\xi \mid \mathcal{F}^\eta\} = \mathbf{M}\{\xi \mid \eta\}$, *где* $\mathcal{F}^\eta = \sigma\{\eta\}$, *называется* ***условным математическим ожиданием СВ*** ξ ***относительно СВ*** $\eta \in \mathbb{R}^n$.

Теорема 4.2.11. *Существует борелевская функция* $g(x)$, $x \in \mathbb{R}^n$ *такая, что* $\mathbf{M}\{\xi \mid \eta\} = g(\eta)$ ($\mathbf{P}$-п.н.).

Пусть ξ — оцениваемая СВ по наблюдениям, образующим случайный вектор $\eta \in \mathbb{R}^n$, а $g(x)$, $x \in \mathbb{R}^n$ — борелевская функция. Тогда будем говорить, что $\widetilde{\xi} = g(\eta)$ — допустимая оценка для ξ по наблюдениям η.

Определение 4.2.23. *Допустимая оценка* $\widehat{\xi} = \widehat{g}(\eta)$ *называется* ***с.к.-оптимальной оценкой*** *для* ξ *по наблюдениям* η, *если для любой допустимой оценки* $\widetilde{\xi} = g(\eta)$ *выполнено*

$$\mathbf{M}\Big\{|\xi - \widehat{\xi}|^2\Big\} \le \mathbf{M}\Big\{|\xi - \widetilde{\xi}|^2\Big\}.$$

Теорема 4.2.12. *Пусть* $\mathbf{M}\big\{|\xi|^2\big\} < \infty$, *тогда с.к.-оптимальная оценка имеет вид*

$$\widehat{\xi} = \mathbf{M}\{\xi \mid \eta\} \quad (\mathbf{P}\text{-п.н.}).$$

Замечание. Предположение о том, что $g(x)$ — борелевская функция, означает, что $\widetilde{\xi} = g(\eta)$ — случайная величина. Утверждение теоремы 4.2.12 имеет важное практическое значение, так как мы получаем общий алгоритм построения с.к.-оптимальной оценки $\widehat{\xi}$ для СВ ξ по наблюдениям η.

Определение 4.2.24. *Пусть* $A \in \mathcal{F}$ *и* $\mathbf{P}\{A\} > 0$. ***Условным математическим ожиданием*** *СВ* ξ ***относительно случайного события*** A *называется СВ*

$$\mathbf{M}\{\xi \mid A\} = \frac{\mathbf{M}\{\xi I_A\}}{\mathbf{P}\{A\}} = \frac{1}{\mathbf{P}\{A\}} \int\limits_A \xi(\omega)\mathbf{P}(d\omega). \tag{4.2.7}$$

Пусть события $H_1, \ldots, H_N$ образуют систему вероятностных гипотез (см. разд. 4.2.1). Если $\mathbf{M}\{\xi\}$ существует, то его можно вычислить по формуле ***полного математического ожидания***:

$$\mathbf{M}\{\xi\} = \sum_{k=1}^{n} \mathbf{P}\{H_k\}\,\mathbf{M}\{\xi \mid H_k\},$$

где $\mathbf{M}\{\xi \mid H_k\}$ вычисляется по формуле (4.2.7).

В заключение раздела рассмотрим практический способ вычисления функции $\mathbf{M}\{\xi \mid \eta = y\} = \widehat{g}(y)$, $y \in \mathbb{R}^n$, где $\xi \in \mathbb{R}^m$, $\eta \in \mathbb{R}^n$. Пусть существует совместная плотность распределения $p_{\xi,\eta}(x, y)$ случайных векторов ξ и η. Тогда

$$\widehat{g}(y) = \int\limits_{\mathbb{R}^m} x\, p_{\xi|\eta}(x \mid y)\, dx, \tag{4.2.8}$$

где $p_{\xi|\eta}(x \mid y) = p_{\xi,\eta}(x, y)/p_\eta(y)$ — ***условная плотность СВ*** ξ ***при условии, что*** $\eta = y$, $p_\eta(y) = \int\limits_{\mathbb{R}^n} p_{\xi,\eta}(x, y) dy$ — n-мерная плотность СВ η, причем $p_\eta(y) \neq 0$. Если же $p_\eta(y) = 0$, то полагают, что $p_{\xi|\eta}(x \mid y) = 0$. Формула (4.2.8) позволяет вычислить реализацию оценки $\widehat{\xi} = \widehat{g}(\eta)$ при условии, что имеется реализация $y \in \mathbb{R}^n$ случайного вектора наблюдений $\eta \in \mathbb{R}^n$.

4.2.6. Гауссовские случайные величины и векторы

Определение 4.2.25. *Функция*

$$\varPhi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-y^2/2} dy, \quad x \in \mathbb{R}^1$$

называется ***интегралом вероятностей*** *или* ***функцией Лапласа***.

Определение 4.2.26. *Случайная величина* $\xi \in \mathbb{R}^1$ *называется* ***гауссовской*** *или* ***нормальной*** *с параметрами* $(m; \sigma^2)$, *где* $\sigma > 0$, *если*

$$F_\xi(x) = \mathbf{P}\{\xi \leq x\} = \varPhi\left(\frac{x - m}{\sigma}\right). \tag{4.2.9}$$

Так как функция Лапласа $\varPhi(x)$ непрерывно дифференцируема на $\mathbb{R}^1$, распределение $F_\xi(x)$ гауссовской СВ имеет плотность

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x - m)^2}{2\sigma^2}\right\}, \quad \sigma > 0.$$

Из определения 4.2.26 следует, что

$$\mathbf{M}\{\xi\} = m, \quad \mathbf{D}\{\xi\} = \sigma^2.$$

Для обозначения гауссовской СВ будем писать $\xi \sim \mathcal{N}(m;\sigma^2)$. Вероятность попадания ξ в произвольный интервал $(\alpha, \beta) \subseteq \mathbb{R}^1$ можно вычислить по следующей известной формуле:

$$\mathbf{P}\{\alpha < \xi < \beta\} = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right).$$

З а м е ч а н и е. Свойство гауссовости распределения сохраняется при линейном преобразовании СВ ξ. Пусть $\xi \sim \mathcal{N}(m_\xi; D_\xi)$, а $\eta = \alpha\xi + \beta$, где $\alpha, \beta \in \mathbb{R}^1$, тогда $\eta \sim \mathcal{N}(m_\eta; D_\eta)$, где $m_\eta = \alpha m_\xi + \beta$; $D_\eta = \alpha^2 D_\xi$.

Для описания ***гауссовского случайного вектора*** (т. е. упорядоченной системы гауссовских СВ) удобно воспользоваться аппаратом характеристических функций.

Пусть $\xi = \{\xi_1, \ldots, \xi_n\}^*$ — вещественный случайный вектор с математическим ожиданием $m_\xi = \{m_{\xi_1}, \ldots, m_{\xi_n}\}^*$ и ковариационной матрицей $R_\xi = \{\mathbf{cov}\{\xi_i, \xi_j\}\}_{i,j=1,\ldots n}$. Пусть также $x = \{x_1, \ldots, x_n\}^* \in \mathbb{R}^n$, $F_\xi(x)$ — n-мерная функция распределения СВ ξ, а i — мнимая единица, т. е. $i^2 = -1$.

О п р е д е л е н и е 4.2.27. *Комплексная функция $\Psi_\xi(\lambda)$, $\lambda \in \mathbb{R}^n$*

$$\Psi_\xi(\lambda) = \mathbf{M}\big\{e^{i\lambda^*\xi}\big\} = \int\limits_{\mathbb{R}^n} e^{i\lambda^* x} dF_\xi(x),$$

называется ***характеристической функцией*** *распределения $F_\xi(x)$.*

Т е о р е м а 4.2.13. *Характеристическая функция однозначно определяет функцию распределения, т. е. если СВ ξ и СВ η имеют одну характеристическую функцию $\Psi_\xi(\lambda) = \Psi_\eta(\lambda)$, $\lambda \in \mathbb{R}^n$, то $F_\xi(x) = F_\eta(x)$, $x \in \mathbb{R}^n$.*

Теперь мы можем ввести понятие ***гауссовского случайного вектора***.

О п р е д е л е н и е 4.2.28. *Случайный вектор $\xi \in \mathbb{R}^n$ имеет* ***n-мерное гауссовское распределение*** *с параметрами $(m_\xi; R_\xi)$, если его характеристическая функция имеет вид*

$$\Psi_\xi(\lambda) = \exp\left\{i\lambda^* m_\xi - \frac{1}{2}\lambda^* R_\xi \lambda\right\}, \quad \lambda \in \mathbb{R}^n,$$

где m_ξ — математическое ожидание, а R_ξ — ковариационная матрица.

З а м е ч а н и я: 1. Нетрудно проверить, что любая компонента ξ_k гауссовского вектора ξ имеет распределение $\mathcal{N}(m_k; D_k)$, где m_k — k-й элемент вектора m_ξ, а D_k — k-й диагональный элемент матрицы R_ξ.

2. Если матрица $R_\xi > 0$ (т. е. положительно определена), то $F_\xi(x)$ имеет плотность распределения

$$p_\xi(x) = (2\pi)^{-n/2}(\det[R_\xi])^{-1/2}\exp\left\{-\frac{1}{2}(x-m_\xi)^* R_\xi^{-1}(x-m_\xi)\right\}, \quad x \in \mathbb{R}^n,$$

где $\det[R_\xi] > 0$ — определитель матрицы R_ξ.

Гауссовские векторы обладают серией замечательных свойств, важнейшие из которых перечислены ниже.

1. Если $\mathbf{cov}\{\xi,\eta\} = \mathbf{M}\{\xi\eta^*\} - m_\xi m_\eta^* = 0$, а вектор $\gamma = \{\xi^*,\eta^*\}^*$ — гауссовский, то ξ и η — независимы.

2. Пусть $\xi \sim \mathcal{N}(m_\xi; R_\xi)$, а $\eta = A\xi + b$, где $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^n$, тогда $\eta \sim \mathcal{N}(m_\eta; R_\eta)$, где

$$m_\eta = Am_\xi + b; \qquad R_\eta = AR_\xi A^*.$$

3. Пусть $\{\xi_n\}$ — последовательность гауссовских случайных векторов. Если $\xi_n \xrightarrow{\text{с.к.}} \xi$, $n \to \infty$, то $\xi \sim \mathcal{N}(m_\xi; R_\xi)$, где $m_\xi = \lim\limits_{n\to\infty}\mathbf{M}\{\xi_m\}$, а $R_\xi = \lim\limits_{n\to\infty}\mathbf{cov}\{\xi_n,\xi_n\}$, причем указанные пределы существуют и конечны.

4. Если $\{\xi_n\}$ — последовательность гауссовских СВ и $\xi_n \xrightarrow{\text{п.н.}} \xi$, $n \to \infty$, то $\xi_n \xrightarrow{\text{с.к.}} \xi$, $n \to \infty$ (в общем случае это не верно!).

5. Теорема 4.2.14 (о нормальной корреляции). *Пусть $\gamma = \{\xi^*,\eta^*\}^*$ — гауссовский вектор такой, что $R_\eta > 0$, тогда*

а) *условное математическое ожидание имеет вид*

$$\widehat{\xi} = \mathbf{M}\{\xi \mid \eta\} = m_\xi + R_{\xi\eta}R_\eta^{-1}(\eta - m_\eta); \tag{4.2.10}$$

б) $\xi \perp \xi - \widehat{\xi}$, *т. е.* ξ *и* $\xi - \widehat{\xi}$ — *независимы;*

в) *пусть* $\Delta\xi = \xi - \widehat{\xi}$, *тогда*

$$\mathbf{M}\{\Delta\xi\} = 0, \quad \mathbf{cov}\{\Delta\xi,\Delta\xi\} = R_\xi - R_{\xi\eta}R_\eta^{-1}R_{\xi\eta}^*;$$

г) *условное математическое ожидание* $\widehat{\xi}$ *имеет гауссовское распределение с параметрами* $(m_\xi; R_{\xi\eta}R_\eta^{-1}R_{\xi\eta}^*)$, *где*

$$R_\xi = \mathbf{cov}\{\xi,\xi\}, \quad R_{\xi\eta} = \mathbf{cov}\{\xi,\eta\}, \quad R_\eta = \mathbf{cov}\{\eta,\eta\}.$$

З а м е ч а н и е. Теорема о нормальной корреляции дает явный вид (4.2.10) с.к.-оптимальной оценки $\widehat{\xi} = \mathbf{M}\{\xi \mid \eta\}$ для ξ по наблюдениям η в гауссовском случае. Заметим, что $\widehat{\xi}$ ***линейно*** зависит от η.

6. Если $\{\gamma, \xi, \eta\}$ составляют гауссовский вектор, а ξ и η — некоррелированные, то

$$\mathbf{M}\{\gamma \mid \xi, \eta\} = \mathbf{M}\{\gamma \mid \xi\} + \mathbf{M}\{\gamma \mid \eta\} - m_\gamma.$$

7. Если компоненты вектора $\xi = \{\xi_1, \ldots, \xi_n\}^*$ — гауссовские и независимые в совокупности, то ξ — гауссовский случайный вектор.

4.2.7. Гильбертово пространство случайных величин с конечным вторым моментом

Для вероятностных приложений наиболее важным является пространство $\mathcal{H}$ случайных величин ξ, определенных на одном и том же вероятностном пространстве $\{\Omega, \mathcal{F}, \mathbf{P}\}$, центрированных и имеющих конечный второй момент:

$$\mathbf{M}\{\xi\} = 0 \quad \text{и} \quad \int\limits_{\Omega} |\xi(\omega)|^2 \mathbf{P}(d\omega) = \mathbf{M}\big\{|\xi|^2\big\} < \infty.$$

Если $\xi, \eta \in \mathcal{H}$, то положим

$$(\xi, \eta) = \mathbf{M}\{\xi\overline{\eta}\} = \mathbf{cov}\{\xi, \eta\}\,.$$

Если $\xi, \eta, \zeta \in \mathcal{H}$, то справедливы следующие свойства операции $(\cdot, \cdot)$:

1) $(\xi, \xi) \geq 0$; если $(\xi, \xi) = 0$, то $\xi = 0$ ($\mathbf{P}$-п.н.);

2) $(a\xi + b\eta, \zeta) = a(\xi, \zeta) + b(\eta, \zeta)$ для любых $a, b \in \mathbb{C}$;

3) $(\eta, \xi) = \overline{(\xi, \eta)}$.

Тем самым $(\cdot, \cdot)$ является ***скалярным произведением*** и определяет ***норму*** в пространстве $\mathcal{H}$:

$$\|\xi\| = (\xi, \xi)^{1/2}.$$

Если $\xi_n \to \xi$, $n \to \infty$ в $\mathcal{H}$, то это означает, что $\mathbf{M}\big\{|\xi_n - \xi|^2\big\} \to 0$. Таким образом, сходимость в $\mathcal{H}$ означает с.к.-сходимость. В силу свойств с.к.-сходимости заключаем, что если $\{\xi_n\}$ фундаментальна, то она сходится к некоторой СВ ξ, причем $\|\xi\|^2 = \mathbf{M}\big\{|\xi|^2\big\} < \infty$, т. е. $\xi \in \mathcal{H}$. Последнее означает, что

$\mathcal{H}$ — пространство со скалярным произведением, полное относительно сходимости по норме, порожденной этим произведением. Итак, $\mathcal{H}$ — ***гильбертово пространство***.

Если $\xi, \eta \in \mathcal{H}$ и $(\xi, \eta) = 0$, то эти СВ называются ***ортогональными***, что обозначается как $\xi \perp \eta$. Понятие ортогональности играет важную роль в задачах оценивания случайных величин.

Пусть $\xi, \eta_1, \ldots \eta_n \in \mathcal{H}$. Предположим, мы хотим оценить случайную величину ξ по наблюдениям $\{\eta_1, \ldots, \eta_n\}$. Как следует из теоремы 4.2.12, такой оценкой является $\mathbf{M}\{\xi \mid \eta_1, \ldots, \eta_n\}$. Однако вычисление условного математического ожидания требует знания совместного закона распределения всех СВ $\{\xi, \eta_1, \ldots, \eta_n\}$, что довольно редко выполняется на практике. Если доступная информация ограничена лишь первыми двумя моментами совместного распределения и оно не является гауссовским, то можно определить ***наилучшую в среднеквадратическом смысле линейную оценку*** для ξ, т. е. ***с.к.-оптимальную линейную оценку***.

Определение 4.2.29. *Случайная величина*

$$\widehat{\xi} = \widehat{l}(\eta_1, \ldots, \eta_n) = \sum_{k=1}^{n} a_k \eta_k$$

называется ***с.к.-оптимальной линейной оценкой*** *для ξ по наблюдениям $\{\eta_1, \ldots, \eta_n\}$, если для любой линейной функции $l(y_1, \ldots, y_n)$ имеет место неравенство*

$$\mathbf{M}\Big\{|\xi - \widehat{\xi}|^2\Big\} \leq \mathbf{M}\big\{|\xi - l(\eta_1, \ldots, \eta_n)|^2\big\}\,.$$

Теорема 4.2.15. *Пусть $\xi, \eta \in \mathcal{H}$ и матрица $R_\eta = \mathbf{cov}\{\eta, \eta\}$ положительно определена. Тогда с.к.-оптимальная линейная оценка $\widehat{\xi}$ вычисляется по формуле*

$$\widehat{\xi} = R_{\xi\eta} R_\eta^{-1} \eta, \tag{4.2.11}$$

где $R_{\xi\eta} = \mathbf{cov}\{\xi, \eta\}$. При этом

$$\|\xi - \widehat{\xi}\|^2 = \mathbf{M}\Big\{|\xi - \widehat{\xi}|^2\Big\} = D_\xi - R_{\xi\eta} R_\eta^{-1} R_{\xi\eta}^*. \tag{4.2.12}$$

Замечания: 1. Формулы (4.2.11), (4.2.12) идентичны формулам теоремы о нормальной корреляции, что не случайно. Действительно, для любой конечной совокупности случайных величин из $\mathcal{H}$ существует семейство гауссовских случайных величин, имеющих те же моменты первого и второго порядка. Для этих гауссовских величин с.к.-оптимальная оценка для ξ есть условное математическое ожидание, которое по теореме о нормальной корреляции является линейной функцией от наблюдений. Поэтому параметры, определяющие эту линейную функцию, одновременно определяют и наилучшую линейную оценку в $\mathcal{H}$.

2. Заметим, что для с.к.-оптимальной линейной оценки $\widehat{\xi}$ также справедливо $\widehat{\xi} \perp \xi - \widehat{\xi}$.

3. Если через $\mathcal{H}(\eta)$ обозначить замкнутое линейное подпространство, порожденное случайным вектором $\eta = \{\eta_1^*, \ldots, \eta_n^*\}^*$, а через $\pi_{\mathcal{H}(\eta)}(\cdot)$ — оператор ортогонального проектирования на $\mathcal{H}(\eta)$ (см. разд. 4.1.8), то очевидно, что с.к.-оптимальная линейная оценка

$$\widehat{\xi} = \pi_{\mathcal{H}(\eta)}(\xi).$$

Оператор $\pi_{\mathcal{H}(\eta)}$ называют также оператором ***условного математического ожидания в широком смысле*** и обозначают $\widehat{\mathbf{M}}\{\cdot \mid \eta\}$. Таким образом, $\widehat{\xi} = \widehat{\mathbf{M}}\{\xi \mid \eta\}$ — с.к.-оптимальная линейная оценка.

4.2.8. Ортогональная стохастическая мера

Пусть задано вероятностное пространство $\{\Omega, \mathcal{F}, \mathbf{P}\}$ и некоторое множество $E \subseteq \mathbb{R}^1$ с алгеброй $\mathcal{E}_0$ его подмножеств. Пусть также $\mathcal{E} = \sigma\{\mathcal{E}_0\}$ — минимальная σ-алгебра, содержащая $\mathcal{E}_0$.

Определение 4.2.30. *Комплексная функция $Z_0(\Delta) = Z_0(\omega; \Delta)$, где $\omega \in \Omega$, $\Delta \in \mathcal{E}_0$, называется **элементарной стохастической мерой**, если $\mathbf{M}\{Z_0(\Delta)\} = 0$, причем выполнены свойства:*

1) $\mathbf{M}\big\{|Z_0(\Delta)|^2\big\} < \infty$ *для всех* $\Delta \in \mathcal{E}_0$;

2) *если* $\Delta_1 \cap \Delta_2 = \varnothing$, *где* $\Delta_1, \Delta_2 \in \mathcal{E}_0$, *то*

$$Z_0(\Delta_1 \cap \Delta_2) = Z_0(\Delta_1) + Z_0(\Delta_2) \quad (\mathbf{P}\text{-п.н.});$$

3) *если* $\{\Delta_n\}$ — *последовательность множеств из* $\mathcal{E}_0$ *таких, что* $\Delta_1 \supseteq \Delta_2 \supseteq \ldots \supseteq \Delta_n \supseteq \ldots$ *и* $\bigcap\limits_{n=1}^{\infty} \Delta_n = \varnothing$ *(т. е.* $\Delta_n \downarrow \varnothing$ *при* $n \to \infty$*), то* $\mathbf{M}\big\{|Z_0(\Delta_n)|^2\big\} \to 0$ *при* $n \to \infty$.

Среди всех элементарных стохастических мер особый интерес для наших целей представляет ортогональная стохастическая мера.

Определение 4.2.31. *Элементарная стохастическая мера* $Z_0(\Delta)$, $\Delta \in \mathcal{E}_0$ *называется* ***ортогональной***, *если для любых непересекающихся* $\Delta_1, \Delta_2 \in \mathcal{E}_0$ *выполнено* $Z_0(\Delta_1) \perp Z_0(\Delta_2)$, *т. е.* $\mathbf{M}\Big\{Z_0(\Delta_1)\overline{Z_0(\Delta_2)}\Big\} = 0$.

Стохастическая мера $Z_0(\cdot)$ тесно связана с некоторой неслучайной мерой, которая определена на $\mathcal{E} = \sigma\{\mathcal{E}_0\}$ и вводится следующим образом. Пусть

$$m_0(\Delta) = \mathbf{M}\big\{|Z_0(\Delta)|^2\big\}, \quad \Delta \in \mathcal{E}_0.$$

Нетрудно проверить, что $m_0(\Delta)$ — конечная мера на $\mathcal{E}_0$. Тогда по теореме Каратеодори (см. разд. 4.1.2) ее можно единственным образом продолжить до меры m, определенной на $\mathcal{E}$.

Определение 4.2.32. *Конечная мера* $m(\Delta)$, $\Delta \in \mathcal{E}$ *называется* ***структурной функцией*** *элементарной ортогональной стохастической меры* $Z_0(\Delta)$, $\Delta \in \mathcal{E}_0$.

Оказывается, меру $Z_0(\cdot)$ можно единственным ($\mathbf{P}$-п.н.) образом продолжить до ортогональной стохастической меры $Z(\cdot)$, определенной на $\mathcal{E}$.

Теорема 4.2.16. *Пусть* $Z_0(\Delta)$, $\Delta \in \mathcal{E}_0$ — *ортогональная элементарная стохастическая мера. Тогда существует единственная* ($\mathbf{P}$-п.н.) *ортогональная мера* $Z(\Delta)$, $\Delta \in \mathcal{E}$ *такая, что для любого* $\Delta \in \mathcal{E}_0$ *выполнено* $Z_0(\Delta) = Z(\Delta)$ ($\mathbf{P}$-п.н.), *причем*

$$\mathbf{M}\big\{|Z(\Delta)|^2\big\} = m(\Delta), \quad \Delta \in \mathcal{E},$$

где $m(\cdot)$ — *структурная функция меры* $Z_0(\cdot)$.

Замечания: 1. Мера $Z(\cdot)$ является σ-аддитивной в с.к.-смысле, т. е. если $\Delta_n \in \mathcal{E}$, $n = 1, 2, \ldots$, $\Delta_m \cap \Delta_k = \varnothing$ при $m \neq k$ и $\Delta = \sum_{n=1}^{\infty} \Delta_n \in \mathcal{E}$, то

$$\mathbf{M}\left\{\left|Z(\Delta) - \sum_{n=1}^{N} Z(\Delta_n)\right|^2\right\} \to 0 \quad \text{при} \quad N \to \infty.$$

2. Меру $m(\Delta)$ также будем называть структурной функцией ортогональной стохастической меры $Z(\cdot)$ на $\mathcal{E}$, так как для любого $\Delta \in \mathcal{E}$ следует

$$m(\Delta) = \mathbf{M}\big\{|Z(\Delta)|^2\big\}.$$

4.2.9. Стохастический интеграл по ортогональной мере

Пусть ортогональная стохастическая мера $Z(\cdot)$ и ее структурная функция $m(\cdot)$ заданы на $\mathcal{E}$. Рассмотрим следующие два гильбертовых пространства: