Содержание

Вопросы и комментарии можно отправлять по e-mail himer2001@mail.ru или бросать в ICQ 257457884.

1. Введение.

1.1. Некоторые сведения из теории множеств.

Определение: *множество* – аксиоматическое понятие, которое может быть интерпретировано как некоторый набор элементов; *пустое множество* – множество, не содержащее элементов. $x \in A$ обозначает, что элемент x принадлежит множеству A.

Определение: множество B называется *подмножеством* А$(A \subseteq B)$, если $\forall\, x \in B\ x \in A$; очевидно, что $\forall\ A\ \varnothing \subseteq A$. Множества А и В *равны*, если $A \subseteq B$ и $B \subseteq A$.

Операции над множествами:

1) *Объединение* – $A \bigcup B = \{x \mid x \in A \bigvee x \in B\}$ (элемент x принадлежит хотя бы одному из множеств); $\bigcup\limits_{\alpha \in A} A_\alpha$ - множество, состоящее из элементов всех $A_\alpha (\alpha \in A)$.

2) *Пересечение* – $A \bigcap B = \{x \mid x \in A, x \in B\}$; $\bigcap\limits_{\alpha \in A} A_\alpha = \{x \mid x \in A_\alpha\ \forall\ \alpha \in A\}$.

3) *Разность* – $A \backslash B = \{x \mid x \in A, x \notin B\}$.

4) *Дополнение* – если $A \subseteq \Omega$, то $\overline{A} = \{x \mid x \in \Omega,\ x \notin A\} = \Omega \backslash A$.

Замечание: в математическом анализе черта над буквой, обозначающей множество, обычно символизирует замыкание A (то есть объединение множества со своей границей); тем не менее, в теории вероятностей этим знаком обозначают дополнение A.

Теорема 1 (законы Моргана): $\overline{\bigcup\limits_{\alpha \in A} A_\alpha} = \bigcap\limits_{\alpha \in A} \overline{A_\alpha}$; $\overline{\bigcap\limits_{\alpha \in A} A_\alpha} = \bigcup\limits_{\alpha \in A} \overline{A_\alpha}$

$\triangle\ \forall\, x \in \overline{\bigcup\limits_{\alpha \in A} A_\alpha}\ \ x \notin A_\alpha\, \forall\, \alpha \in A \Rightarrow \forall\, \alpha \in A\ x \in \overline{A_\alpha} \Rightarrow x \in \bigcap\limits_{\alpha \in A} \overline{A_\alpha} \Rightarrow \overline{\bigcup\limits_{\alpha \in A} A_\alpha} \subseteq \bigcap\limits_{\alpha \in A} \overline{A_\alpha}$. $\forall\, x \in \bigcap\limits_{\alpha \in A} \overline{A_\alpha}\ \ x \in \overline{A}_\alpha\ \forall\ \alpha \in A \Rightarrow \forall\ \alpha \in A\ x \notin A_\alpha \Rightarrow x \notin \bigcup\limits_{\alpha \in A} A_\alpha \Rightarrow x \in \overline{\bigcup\limits_{\alpha \in A} A_\alpha} \Rightarrow \bigcap\limits_{\alpha \in A} \overline{A_\alpha} \subseteq \overline{\bigcup\limits_{\alpha \in A} A_\alpha}$. Таким образом, $\overline{\bigcup\limits_{\alpha \in A} A_\alpha} = \bigcap\limits_{\alpha \in A} \overline{A_\alpha}$. Второе равенство доказывается аналогично. ■

Определение: *отображением* f множества A на B называется сопоставление каждому элементу $a \in A$ одного элемента $b \in B$, называемого *образом* $a:\ f(a) = b$. Всякий элемент $a \in A:\ f(a) = b$ называется *прообразом* b в A. Если $C \subseteq A$, то $f(C) = \{f(a),\ a \in C\}$ называется *образом* C в B; если $D \subseteq B$, то $f^{-1}(D) = \{a \in A:\ \ f(a) \in D\}$ называется *полным прообразом* D в A.

Определение: множества A и B *эквивалентны*(равномощны), если существует взаимно однозначное отображение A на B (то есть каждому $a \in A$ соответствует единственное $b \in B : b = f(a)$ и $\forall\ b\ \exists!\ a = f^{-1}(b)$). Таким образом, все множества могут разбиты на *классы эквивалентности* (то есть классы одинаковой мощности).

Определение: множество A называется *бесконечным*, если оно эквивалентно какому-либо своему нетривиальному (то есть не совпадающему с A) подмножеству и *конечным* в противном случае. *Мощность* $(\overline{\overline{A}})$ является числом, характеризующим тот или иной класс эквивалентности; для конечных множеств мощность равна числу элементов. Для бесконечных множеств мощность характеризует тип множества; например, все множества, эквивалентные $\mathbb{N}$, называют *счётными*, а множества, эквивалентные [0,1] - *континуальными*. Можно также показать (см. мат. анализ, 1.1.2), что конечное или счётное объединение счётных множеств счётно, а $\mathbb{R}$ континуально.

Определение: мощности множеств A и B совпадают$(\overline{\overline{A}} = \overline{\overline{B}})$, если А и В эквивалентны и $\overline{\overline{A}} < \overline{\overline{B}}$, если A эквивалентно какому-либо нетривиальному подмножеству B.

Определение: $2^A = \{C \mid C \subseteq A\}$ называется *булеаном* (множеством всех подмножеств) A. Если A конечно $(\overline{\overline{A}} = n \in \mathbb{N})$, то $\overline{\overline{2^A}} = 2^n$ (число подмножеств A определяется как число двоичных векторов длины n, в которых 1 соответствует наличию элемента в подмножестве, а 0 – его отсутствию, см. 1.2).

Определение: Ω – произвольное непустое множество; $\mathfrak{F}$ – набор подмножеств Ω – называется *алгеброй*, если $\forall\ A,\ B \in \mathfrak{F}\ \ \overline{A} \in \mathfrak{F},\ A \bigcup B \in \mathfrak{F}$. $\mathfrak{F}$ называется *σ-алгеброй*, если $\forall\ A_1, \dots A_n, \dots \in \mathfrak{F}\ \ \overline{A}_i \in \mathfrak{F}\ \forall\ i \in \mathbb{N},\ \bigcup\limits_{n\in\mathbb{N}} A_n \in \mathfrak{F}$.

Замечание: из определений следует ряд свойств алгебр и σ-алгебр; если $\mathfrak{F}$ - алгебра, то $\forall\ A,\ B \in \mathfrak{F}\ A \bigcap B = \overline{\overline{A} \bigcup \overline{B}} \in \mathfrak{F},\ \Omega = A \bigcup \overline{A} \in \mathfrak{F},\ \varnothing = A \bigcap \overline{A} \in \mathfrak{F}$. Если $\mathfrak{F}$ - σ-алгебра, то, по теореме 1, $\bigcap\limits_{n\in\mathbb{N}} A_n = \overline{\bigcup\limits_{n\in\mathbb{N}} \overline{A_n}} \in \mathfrak{F}$. Поэтому $\varnothing = A_1 \bigcap \overline{A_1} \bigcap A_1 \bigcap \dots \in \mathfrak{F} \Rightarrow A \bigcup B = A \bigcup B \bigcup \varnothing \bigcup \dots \in \mathfrak{F} \Rightarrow \Omega = A \bigcup \overline{A} \in \mathfrak{F}$.

1.2. Некоторые сведения из комбинаторики.

Комбинаторика - раздел математики, занимающийся подсчётом числа комбинаций элементов конечного множества, составленных при определённых условиях.

Определение *генеральной совокупностью* объёма n называется произвольное конечное множество $E = \{a_1, \dots a_n\}$; выборкой их генеральной совокупности объёма r называется произвольный набор элементов E: $A = \{a_{i_1}, \dots a_{i_r}\}$.

Теорема 1 (правило умножения): если существует N_1 способов выполнить действие 1, а затем N_2 способов выполнить действие 2, то существует $N_1 N_2$ способов выполнить 1 и 2 последовательно.

Теорема 2 (правило сложения): если существует N_1 способов выполнить действие 1 и N_2 способов выполнить действие 2, то существует $N_1 + N_2$ способов выполнить одно из действий - 1 или 2.

$\triangle$ Данные утверждения вполне очевидны, поэтому их доказательства не приводятся.

Определение: выборка называется *упорядоченной*, если имеет значение порядок её элементов, и *неупорядоченной* в обратном случае. Выборка проводится *без возвращения*, если каждый элемент генеральной совокупности входит в неё не более одного раза, и *с возвращением* в обратном случае.

Определение: *двоичный вектор* длины n - вектор длины n, компонентами которого являются нули и единицы.

Подсчёт числа выборок (N) объёма r из генеральной совокупности объёма n:

Выборки	Упорядоченные	Неупорядоченные
С возвращением	n^r	C^r_{n+r-1}
Без возвращения	A^r_n	C^r_n

1) *Упорядоченные с возвращением*: в выборке на каждом месте может находиться один из n элементов, поэтому, по правила умножения, $N = n \cdot \ldots \cdot n = n^r$. В частности, имеется 2^n двоичных векторов длины n.

2) *Упорядоченные без возвращения*: на первом месте выборки может находиться один из n элементов, на втором – один из $n-1$ элементов, и так далее. $N = n(n-1) \cdot \ldots \cdot (n-r+1) = \dfrac{n!}{(n-r)!} = A^r_n$ – число *размещений.*

3) *Неупорядоченные без возвращения*: конкретное размещение из n по r могут реализовываться (в зависимости от порядка элементов) $r!$ способами (на первом месте - любой из r элементов, на втором – любой из $r-1$ элементов и так далее), поэтому $N = \frac{A^r_n}{r!} = \frac{n!}{r! \cdot (n-r)!} = C^r_n$ – число *сочетаний.*

4) *Неупорядоченные с возвращением*: сопоставим каждой выборке двоичный вектор, в который входят единицы по числу раз, которое элемент данного типа входит в выборку, и нули, разделяющие элементы (например, если $E = \{a_1, a_2, a_3\}$ и $A = \{a_1, a_2, a_1, a_1\}$, то вектор имеет вид (111010)). Таким образом получим двоичный вектор длины $n + r - 1$, содержащий r единиц; всего существует $N = C^r_{n+r-1}$ таких векторов.

Подсчёт числа (N) размещений n частиц по r ячейкам:

Частицы	Различимые	Неразличимые
С запретом	$\dfrac{n!}{n_1! \cdot \ldots \cdot n_r!}$	C_{n-1}^{r-1}
Без запрета	r^n	C_{n+r-1}^{r-1}

1) *Различимые частицы без запрета*: каждая частица может попасть в любую из r ячеек, поэтому $N = r^n$.

2) *Различимые частицы, с запретом* (в i-ую ячейку попадает n_i элементов, $\sum\limits_{i=1}^{r} n_i = n$): в первую ячейку могут попасть n_1 из n элементов, во вторую n_2 из $n - n_1$ элементов, и так далее. Поэтому $N = C_n^{n_1} C_{n-n_1}^{n_2} \cdot \ldots \cdot C_{n-n_1-\ldots-n_{r-1}}^{n_r} = \dfrac{n!}{(n-n_1)!n_1!} \cdot \dfrac{(n-n_1)!}{(n-n_1-n_2)!n_2!} \cdot \ldots \cdot \dfrac{(n-n_1-\ldots-n_{r-1})!}{0!n_r!} = \dfrac{n!}{n_1! \cdot \ldots \cdot n_r!}$.

3) *Неразличимые частицы, без запрета*: сопоставим каждому размещещнию двоичный вектор, содержащий единицы по числу частиц в той или иной ячейке и нули как "разделители" между ячейками. Так получится двоичный вектор длины $n + r - 1$, содержащий n единиц, то есть $N = C_{n+r-1}^{n} = C_{n+r-1}^{r-1}$.

4) *Неразличимые частицы, с запретом* (ни одна ячейка не остаётся пустой): снимем запрет, заранее поместив по одной частице в каждую ячейку; тогда, согласно п. 3,

$$N = C_{n-r+r-1}^{r-1} = C_{n-1}^{r-1}.$$

2. Аксиоматика теории вероятностей и её простейшие применения.

2.1. Аксиомы теории вероятностей.

Определение: *случайное событие* – событие, которое может произойти с различными, взаимно исключающими исходами; в теории вероятностей рассматриваются только те случайные события, для которых возможна *повторяемость* (возможность многократного проведения эксперимента в одних и тех же условиях), а также наблюдается *статистическая устойчивость частот* (*частота* – отношение числа выпадений того или иного исхода к общему числу испытаний). Подобные события называют *стохастическими.*

Определение: *пространство элементарных исходов* Ω является аксиоматическим понятием и представляет собой непустое множество, элементы которого символизируют тот или иной исход рассматриваемого процесса. Элементы Ω называются *элементарными событиями* (*элементарными исходами*), а подмножества Ω – *событиями.* Говорят, что событие A произошло, если реализовался хотя бы один из входящих в него элементарных исходов.

Определение: A, B – события; *суммой событий* $(A+B)$ называется событие, состоящее в том, что одно из событий – A или B – произошло. *Произведением событий* (AB) называется событие, состоящее в том, что произошли оба события – и A, и B. *Разностью событий* $(A\backslash B)$ называется событие, состоящее в том, что A произошло, а B – нет. Событием, *обратным* к A $(\overline{A})$ называется событие, состоящее в том, что A не произошло. Очевидно, $A+B=A\bigcup B,\ AB=A\bigcap B.$

Определение: $\mathfrak{F}$ – σ-алгебра событий. P:$\mathfrak{F} \longrightarrow \mathbb{R}$ называется вероятностной мерой (или просто *вероятностью*), если она удовлетворяет следующим условиям (*аксиомы вероятности*): $\forall\ A_i \in \mathfrak{F}$ $(i \in \mathbb{N})$

1) $P(A_i) \geq 0,$

2) $P(\Omega) = 1,$

3) Если $A_iA_j = \varnothing\ \forall\ i, j \in \mathbb{N}$ $(i \neq j)$, то $P\left(\sum\limits_{n=1}^{\infty} A_n\right) = \sum\limits_{n=1}^{\infty} P(A_i)$ (здесь $\sum$ понимается как числовой ряд, схоящийся абсолютно).

Определение: *вероятностным пространством* называется набор $(\Omega, \mathfrak{F}, P)$, описывающий данное случайное событие.

Замечание: третья аксиома вероятности (аксиома σ-аддитивности) может быть заменена на две – *аксиомы конечной аддитивности* и *непрерывности*: $\forall\ A, B \in \mathfrak{F}$: $AB = \varnothing$ $P(A+B) = P(A) + P(B)$; $\forall\ A_1, \dots A_n, \dots \in \mathfrak{F}$: $A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \supseteq \dots,\ \prod\limits_{n \in \mathbb{N}} A_n = \varnothing$ $\lim\limits_{n\to\infty} P(A_n) = 0.$

$\triangle$ $\Rightarrow$. Пусть $A_1 \supseteq \dots \supseteq A_n \supseteq \dots,\ \prod\limits_{n\in\mathbb{N}} A_n = \varnothing$. $A_n = \sum\limits_{k=n}^{\infty} A_k\overline{A}_{k+1} + \prod\limits_{k=n}^{\infty} A_k$; тогда, по аксиоме σ-аддитивности, $P(A_n) = \sum\limits_{k=n}^{\infty} P\left(A_k\overline{A}_{k+1}\right) + P\left(\prod\limits_{k=n}^{\infty} A_k\right) = \sum\limits_{k=n}^{\infty} P\left(A_k\overline{A}_{k+1}\right) \to 0,\ n \to \infty$ (как остаток схоящегося ряда $\sum\limits_{k=1}^{\infty} P\left(A_k\overline{A}_{k+1}\right) = P(A_1)$).

$\Leftarrow$. Пусть $A_1, \dots A_n, \dots \in \mathfrak{F}; A_iA_j = \varnothing\ \forall\ i, j \in \mathbb{N}$ $(i \neq j)$; $A = \sum\limits_{n=1}^{\infty} A_n,\ B_n = \sum\limits_{k=n}^{\infty} A_k \Rightarrow$ $B_{n+1} \subseteq B_n$. Если наступило B_n, то наступило только одно из $A_k(k \geq n) - A_m$, то есть $\forall\ l > m\ A_l$ не наступило $\Rightarrow$ $P\left(\prod\limits_{k=m+1}^{\infty}\right) = 0 \Rightarrow P(B_n) \to 0, n \to \infty$; но $P(A) = \sum\limits_{k=1}^{n} A_k + B_{n+1}$, поэтому, в пределе при $n \to \infty$, $P(A) = P\left(\sum\limits_{k=1}^{\infty}\right) = \sum\limits_{k=1}^{\infty} P(A_k)$. $\blacksquare$

Теорема 1 (свойства вероятности): $\forall\ A, B \in \mathfrak{F}$

1) $B \subseteq A \Rightarrow P(B) \leq P(A),\ P(A \backslash B) = P(A) - P(B)$;

2) $0 \leq P(A) \leq 1$;

3) $P(\overline{A}) = 1 - P(A)$;

4) $P(A + B) = P(A) + P(B) - P(AB)$.

$\triangle$ 1) $A = A\overline{B} + B,\ (A\overline{B})B = \varnothing \Rightarrow P(A) = P(A \backslash B) + P(B) \Rightarrow P(A) \geq P(B),\ P(A \backslash B) = P(A) - P(B)$.

2) $\forall\ A \in \mathfrak{F}\ A \subseteq \Omega \Rightarrow P(A) \leq P(\Omega) = 1$.

3) $A + \overline{A} = \Omega \Rightarrow P(A) + P(\overline{A}) = P(\Omega) = 1 \Rightarrow P(\overline{A}) = 1 - P(A)$.

4) $A + B = \overline{A}B + \overline{B}A + AB \Rightarrow P(A + B) = P(A) - P(AB) + P(B) - P(AB) + P(AB) = P(A) + P(B) - P(AB)$. $\blacksquare$

Замечание: данные аксиомы были введены А. Н. Колмогоровым (поэтому их часто называют аксиомами Колмогорова); система этих аксиом непротиворечива (то есть существуют функции, ей удовлетворяющие) и неполна (то есть не задаёт вероятностную меру однозначно).

2.2. Вероятность в дискретных пространствах элементарных исходов.

Определение: пространство элементарных исходов *дискретно*, если оно конечно или счётно; в этом случае $\mathfrak{F}$ выбирается как 2^{Ω}.

Определение: Ω – дискретное пространство элементарных исходов; тогда $P\colon \Omega \to \mathbb{R}$ называется вероятностной мерой (*вероятностью*), если $\forall\ \omega \in \Omega\ P(\omega) \geq 0,\ \sum\limits_{\omega \in \Omega} P(\omega) = 1$. В этом случае $\forall\ A \in \mathfrak{F}\quad P(A) = \sum\limits_{\omega \in A} P(\omega)$ (все $\sum$ понимаются как конечные суммы или сходящиеся числовые ряды).

Теорема 1 (теорема сложения): Ω – дискретное пространство элементарных исходов; $\forall\ A, B \in \mathfrak{F}\colon AB = \varnothing\ P(A + B) = P(A) + P(B)$.

$\triangle$ $P(A + B) = \sum\limits_{\omega \in A+B} P(\omega) = \sum\limits_{\omega \in A} P(\omega) + \sum\limits_{\omega \in B} P(\omega) = P(A) + P(B)$. $\blacksquare$

Следствие: $\forall\ A, B \in \mathfrak{F}\ \ P(A + B) = P(A) + P(B) - P(AB);\ P(A + B) \leq P(A) + P(B)$.

$\triangle$ $A + B = \overline{A}B + \overline{B}A + AB \Rightarrow P(A + B) = \sum\limits_{\omega\colon\ \omega \in B,\ \omega \notin A} P(\omega) + \sum\limits_{\omega\colon\ \omega \in A,\ \omega \notin B} P(\omega) + \sum\limits_{\omega\colon\ \omega \in A,\ \omega \in B} P(\omega) = \sum\limits_{\omega \in A} P(\omega) + \sum\limits_{\omega \in B} P(\omega) - \sum\limits_{\omega\colon\ \omega \in A,\ \omega \in B} P(\omega) = P(A) + P(B) - P(AB)$. Ряды сходятся абсолютно, поэтому перестановки не изменяют их суммы (см. математический анализ, 5.1.4). $\blacksquare$

Теорема 2 (обобщённая теорема сложения): Ω – дискретное пространство элементарных исходов; $\forall\ A_1, \ldots A_n \in \mathfrak{F}\colon A_iA_j = \varnothing\ \forall\ i, j = \overline{1, n}\ (i \neq j)\ P\left(\sum\limits_{k=1}^{n} A_k\right) = \sum\limits_{k=1}^{n} P(A_k)$.

Следствие: $\forall\ A_1, \ldots A_n \in \mathfrak{F}\quad P\left(\sum\limits_{k=1}^{n} A_k\right) = \sum\limits_{k=1}^{n} P(A_k) - \sum\limits_{1 \leq i_1 < i_2 \leq n} P\left(A_{i_1}A_{i_2}\right) + \sum\limits_{1 \leq i_1 < i_2 < i_3 \leq n} P\left(A_{i_1}A_{i_2}A_{i_3}\right) - \ldots + (-1)^{n+1} P(A_1 \ldots A_n);\ P\left(\sum\limits_{k=1}^{n} A_k\right) \leq \sum\limits_{k=1}^{n} P(A_k)$ (очевидно, $P(A_{i_1} \ldots A_{i_m}) \geq \sum\limits_{k=1}^{n} P(A_{i_1} \ldots A_{i_m} A_k)\ \forall\ m = \overline{1, n-1}$).

$\triangle$ Доказательство проводится по индукции с использованием теоремы 1.

Теорема 3 (σ-аддитивности): Ω – дискретное пространство элементарных исходов; $\forall\ A_1, \ldots A_n, \ldots\ \in \mathfrak{F}\colon A_iA_j = \varnothing\ \forall\ i, j = \overline{1, n}\ (i \neq j)\ P\left(\sum\limits_{k=1}^{\infty} A_k\right) = \sum\limits_{k=1}^{\infty} P(A_k)$.

$\triangle$ $P\left(\sum\limits_{k=1}^{\infty} A_k\right) = \sum\limits_{k=1}^{\infty}\sum\limits_{\omega\in A_i} P(\omega) = \sum\limits_{k=1}^{\infty} P(A_k)$. Данный ряд сходится, так как любая его частичная сумма ограничена сверху $\sum\limits_{\omega\in\Omega} P(\omega) = 1$. $\blacksquare$

2.3. Примеры задания вероятности.

Определение: Ω – конечное множество мощности n; тогда $\forall\ A \subseteq \Omega\ P(A) = \frac{\overline{\overline{A}}}{n}$ – *классическое определение вероятности.*

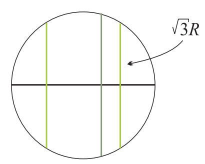
Определение: Ω – измеримая по Жордану (см. математический анализ, 4.1.1) область $\mathbb{R}^n$; для любого A – измеримого по Жордану подмножества Ω – $P(A) = \frac{\mu(A)}{\mu(\Omega)}$: *геометрическая вероятность* (элементарным исходом является попадание в конкретную точку Ω).

Замечание: очевидно, что выбор вероятности с использованием этих схем имеет смысл только в том случае, когда все элементарные исходы равновероятны.

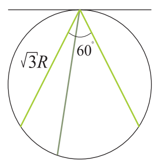
Пример (парадокс Бертрана): требуется определить, с какой вероятностью хорда, проведённая произвольным образом в окружности единичного радиуса, превысит по длине сторону вписанного в эту окружность правильного треугольника (то есть превысит $\sqrt{3}$).

В данном случае есть по крайней мере три способа подсчёта геометрической вероятности:

1) Проведём диаметр, перпендикулярный к хорде, и подсчитаем вероятность, как отношение длины "центрального участка"(того, через который проходят хорды длиннее $\sqrt{3}$) к целому диаметру. $P = \dfrac{\frac{1}{2}+\frac{1}{2}}{2} = \dfrac{1}{2}$.



2) Рассмотрим угол φ между хордой и касательной к окружности – хорда больше $\sqrt{3}$, если $60^0 < \varphi < 120^0$ (то есть хорда лежит внутри вписанного в окружность правильного треугольника). Таким образом, $P = \dfrac{120-60}{180-0} = \dfrac{1}{3}$.



3) Хорда превысит по длине сторону правильного треугольника, если она пересечёт окружность, вписанную в этот треугольник, то есть вероятность рассчитывается как отношение площадей вписанной и описанной окружностей: $P = \frac{\pi\left(\frac{1}{2}\right)^2}{\pi 1^2} = \frac{1}{4}$.

Таким образом, выбор вероятности неоднозначен.

2.4. Условная вероятность.

Определение: $A, B \in \mathfrak{F},\ P(B) > 0$; тогда *условной вероятностью* события A при условии B называется $P_B(A) = P(A|B) = \frac{P(AB)}{P(B)}$.

Теорема 1 (теорема умножения): $A, B \in \mathfrak{F};\ P(A),\ P(B) > 0$; тогда $P(AB) = P(A|B)\cdot$ $\cdot P(B) = P(B|A)P(A)$.

$\triangle$ Непосредственное следует из определения условной вероятности.

Следствие (обобщённая теорема умножения): $A_1, \ldots A_n \in \mathfrak{F};\ \forall\ k = \overline{1, n-1}$ $P(A_1 \ldots A_k) > 0$; тогда $P(A_1 \ldots A_n) = P(A_1)P(A_2|A_1)\cdot\ldots\cdot P(A_n|A_1\ldots A_{n-1})$.

$\triangle$ Доказательство проводится по индукции, исходя из теоремы.

Определение: $A, B \in \mathfrak{F}$; события A и B *независимы*, если $P(AB) = P(A)P(B)$. Очевидно, что в случае $P(A), P(B) > 0$ это определение эквивалентно $P(A|B) = P(A)$, $P(B|A) = P(B)$.

Определение: $A_1, \ldots A_n \in \mathfrak{F}$ $(n \geq 3)$; события $A_1, \ldots A_n$ *независимы в совокупности*, если $\forall\ k = \overline{2, n},\ i_1, \ldots i_k:\ 1 \leq i_1 < \ldots < i_k \leq n\ P\left(A_{i_1}\ldots A_{i_k}\right) = P\left(A_{i_1}\right)\ldots P\left(A_{i_k}\right)$.

Замечание: данное определение независимости событий часто оказывается шире реальной независимости событий, поэтому иногда вопрос о независимости событий решают не математически, а по результатам эксперимента.

Пример (Бернштейна): попарная независимость событий не обязательно означает их независимость в совокупности; рассмотрим правильный тетраэдр, три грани которого окрашены, соответственно, в красный, зелёный и синий цвета, а четвёртая – во все три цвета одновременно (например, грань разбита на три части, окрашенные в разные цвета). Событие 1 заключается в выпадении грани, на которой имеется красный цвет, событие 2 – зелёный цвет, событие 3 – синий цвет. Тогда, согласно классическому определению вероятности, $P_1 = P_2 = P_3 = \frac{2}{4} = \frac{1}{2}$, $P_{12} = P_{13} = P_{23} = \frac{1}{4} = P_1P_2$, однако $P_{123} = \frac{1}{4} \neq P_1P_2P_3$.

Определение: $H_1, \ldots H_n \in \Omega$ образуют *полную группу событий*, если $\bigcup\limits_{i=1}^{n} H_i = \Omega$; $H_i \neq \varnothing$; $H_iH_j = \varnothing \ \forall\ i, j = \overline{1, n} : \ i \neq j$.

Теорема 2 (формула полной вероятности): $H_1, \ldots H_n$ – полная группа событий; тогда $\forall\ A \in \mathfrak{F}\ P(A) = \sum\limits_{i=1}^{n} P(A|H_i)P(H_i)$.

$\triangle$ $A = A\Omega = A(H_1 \bigcup \ldots \bigcup H_n) = AH_1 \bigcup \ldots \bigcup AH_n \Rightarrow (AH_i \bigcap AH_j \neq \varnothing\ \forall\ i, j = \overline{1, n} : i \neq j)$ $P(A) = \sum\limits_{i=1}^{n} P(AH_i) =$ (теорема 1) $= \sum\limits_{i=1}^{n} P(A|H_i)P(H_i)$. $\blacksquare$

Замечание: теорема верна также для случаев $\Omega = \bigcup\limits_{i=1}^{\infty} H_n$, а также в случае $A \subseteq (H_1 \bigcup \ldots \bigcup H_n) \subset \Omega$.

Теорема 3 (формула Байеса): $H_1, \ldots H_n$ – полная группа событий; тогда $\forall\ A \in \mathfrak{F}$, $\forall\ k = \overline{1, n}$ $P(H_k|A) = \dfrac{P(AH_k)}{P(A)} = \dfrac{P(A|H_k)P(H_k)}{\sum\limits_{i=1}^{n} P(A|H_i)P(H_i)}$.

$\triangle$ Формула непосредственно следует из теорем 1 и 2.

2.5. Схема Бернулли.

Определение: *схемой Бернулли* называется последовательность n одинаковых независимых испытаний, имеющих 2 возможных несовместных исхода, достигаемых с вероятностями $p(n)$ (успех) и $q(n) = 1 - p(n)$ (неуспех).

Определение: *полиномиальной схемой* называется последовательность n одинаковых независимых испытаний, имеющих m несовместных исходов, достигаемых с вероятностями $p_1(n), \ldots p_m(n) : \sum\limits_{i} p_i(n) = 1$. Очевидно, что схема Бернулли является биномиальной – частным случаем полиномиальной схемы. Пространством элементарных исходов для полиномиальной схемы является множество всех векторов длины n (число испытаний), компонентами которых являются натуральные числа от 1 до m, соответствующие тому или иному элементарному исходу.

Теорема 1: если μ_n - число успехов в схеме Бернулли, состоящей из n испытаний, то $P_n(m) = P\{\mu_n = m\} = C_n^m p^m q^{n-m}$.

$\triangle$ Число успехов определяется числом единиц в двоичном векторе длины n, описывающем данный результат; вероятность наличия m единиц равна вероятности появления m единиц и $n - m$ нулей, умноженной на число сочетаний из n по m. $\blacksquare$

Следствие: если μ_i – число выпадений i-го исхода в полиномиальной схеме, состоящей из n испытаний, то $P\{\mu_1 = n_1, \ldots \mu_m = n_m\} = \dfrac{n!}{n_1! \ldots n_m!} p_1^{n_1} \ldots p_m^{n_m}$.

$\triangle$ Вероятность равна числу соответствующих сочетаний (n_1 единиц, n_2 двоек, … n_m чисел m – см. 1.2), умноженному на вероятность выпадения n_i раз i-го исхода.

Теорема 2 (Пуассона): p_n – вероятность успеха в схеме Бернулли, состоящей из n испытаний; $\lambda_n = np_n \to \lambda > 0$, $n \to \infty$; тогда для фиксированного числа $m < n$ $P_n(m) \to \dfrac{\lambda^m e^{-\lambda}}{m!}$, $n \to \infty$.

$\triangle$ $P_n(m) = C_n^m p^m q^{n-m} = C_n^m \left(\frac{\lambda_n}{n}\right)^m \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-m} = \frac{n(n-1)\dots(n-m+1)}{m!} \left(\frac{\lambda_n}{n}\right)^m \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-m} \Rightarrow$

$$\lim_{n\to\infty} P_n(m) = \frac{\lambda^m}{m!} \cdot \lim_{n\to\infty} \frac{n(n-1)\dots(n-m+1)}{n^m} \cdot \frac{\left(1-\frac{\lambda_n}{n}\right)^n}{\left(1-\frac{\lambda_n}{n}\right)^m} = \frac{\lambda^m}{m!} \cdot 1 \cdot \lim_{n\to\infty} \exp\left(n\left(1-\frac{\lambda_n}{n}\right)\right) =$$

$$= \frac{\lambda^m}{m!} \cdot \lim_{n\to\infty} \exp(-\lambda_n) = \frac{\lambda^m e^{-\lambda}}{m!}. \blacksquare$$

Замечание: эта теорема может быть использована для приближённого вычисления числа успехов в схеме Бернулли; при больших n и малых p $P_n(m) \approx \dfrac{(np)^m e^{-np}}{m!}$, причём $\left| C_n^m p^m q^{n-m} - \dfrac{(np)^m e^{-np}}{m!} \right| \le np^2$.

Теорема 3 (локальная предельная теорема Муавра-Лапласа – без доказательства): p_n – вероятность успеха в схеме Бернулли, состоящей из n испытаний; $p_n = p = \text{const}$; B – множество таких m, при которых $x_{nm} = \dfrac{m - np}{\sqrt{npq}}$ равномерно ограничены по n; тогда

$$P_n(m) = \frac{1}{\sqrt{2\pi npq}} \cdot \exp\left(-\frac{(m-np)^2}{2npq}\right)\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

Теорема 4 (интегральная предельная теорема Муавра-Лапласа – без доказательства):

$$\forall\, a, b \in \mathbb{R}\ P\left\{a \le \frac{\mu_n - np}{\sqrt{npq}} \le b\right\} \underset{n\to\infty}{\overset{a,b}{\rightrightarrows}} \frac{1}{\sqrt{2\pi}} \cdot \int\limits_a^b e^{-\frac{u^2}{2}} du.$$

Следствие:

$$P\left\{m_1 \le \mu_n \le m_2\right\} \simeq \frac{1}{\sqrt{2\pi}} \int\limits_{x_{m_1}}^{x_{m_2}} e^{-\frac{u^2}{2}} du = \Phi\left(x_{m_2}\right) - \Phi\left(x_{m_1}\right),$$

где $x_{m_i} = \frac{m_i - np}{\sqrt{npq}}$ $(i = 1, 2)$; $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{u^2}{2}} du$ – табличная величина.

Замечание: для приближённых вычислений при $np \le 20$ используют теорему Пуассона и интегральную предельную теорему Муавра-Лапласа при бо́льших значениях np.

3. Случайные величины.

3.1. Одномерные случайные величины и их распределения.

Определение: *борелевской алгеброй* $\mathfrak{B}(\mathbb{R})$ числовых множеств называется набор множеств, полученных применением конечное или счётное число раз теоретико-множественных операций к набору $\{(-\infty, x)\}_{x\in\mathbb{R}}$; элементы такой алгебры называются *борелевскими множествами.* Очевидно, что $\mathfrak{B}(\mathbb{R})$ является σ-алгеброй. Аналогично определяется $\mathfrak{B}(\mathbb{R}^n)$.

Замечание: борелевская алгебра $\mathfrak{B}(\mathbb{R})$ может быть порождена различными числовыми множествами ($(-\infty, x]$, $[x_1, x)$, (x_1, x), $(x, x_1]$, $(x, +\infty)$, $[x, +\infty)$), где $x \in \mathbb{R}$, а x_1 фиксированы.

Определение: $\xi : \Omega \to \mathbb{R}$ – *случайная величина*, если выполнено одно из условий

1) $\forall\ x \in \mathbb{R}\ \ (\xi^{-1}(-\infty, x)) \in \mathfrak{F}$; *функцией распределения* случайной величины называется $F_\xi(x) = P\{\xi < x\} = P\left(\xi^{-1}(-\infty, x)\right)$; или

2) $\forall\ B \in \mathfrak{B}(\mathbb{R})\ \xi^{-1}(B) \in \mathfrak{F}$; тогда $P\{\xi \in B\} = P\left(\xi^{-1}(B)\right)$ – *распределением вероятностей.*

Теорема 1 (без доказательства): функция $F_\xi(x)$ позволяет определить все значения $P\{\xi \in B\}$, то есть сформулированные определения случайной величины эквивалентны.

Теорема 2 (свойства функции распределения): ξ – случайная величина; $F(x) = F_\xi(x)$, тогда

1) $\forall\ x_1, x_2 \in \mathbb{R} :\ x_1 < x_2\ F(x_1) \leq F(x_2)$;

2) $\lim\limits_{x\to-\infty} F(x) = 0,\ \lim\limits_{x\to+\infty} F(x) = 1$;

3) $\forall\ x_0 \in \mathbb{R}\ \lim\limits_{x\to x_0-0} F(x) = F(x_0)$;

4) $F(x)$ имеет конечное или счётное число точек разрыва.

$\triangle$ 1) $(-\infty, x_1) \subset (-\infty, x_2) \Rightarrow \xi^{-1}(-\infty, x_1) \subseteq \xi^{-1}(-\infty, x_2) \Rightarrow P\{\xi < x_1\} \leq P\{\xi < x_2\} \Rightarrow F(x_1) \leq F(x_2)$.

2) $(-\infty, -1) \supset (-\infty, -2) \supset \ldots \supset (-\infty, -n) \supset \ldots,\ \bigcap\limits_{n\in\mathbb{N}}(-\infty, -n) = \varnothing \Rightarrow \xi^{-1}\left(\bigcap\limits_{n\in\mathbb{N}}(-\infty, -n)\right) = \varnothing \Rightarrow$ (аксиома непрерывности) $\lim\limits_{x\to-\infty} F(x) = P(\varnothing) = 0$ (на самом деле, показано существование частичного предела при $x \to -\infty$, однако, поскольку F монотонна, то, по теореме математического анализа (см. 1.3.2), существует и искомый предел). $\lim\limits_{x\to+\infty} F(x) = P\{\xi < +\infty\} = P(\Omega) = 1$.

3) $(-\infty, x_0-1) \supset (-\infty, x_0-\frac{1}{2}) \supset \ldots \supset (-\infty, x_0-\frac{1}{n}) \supset \ldots,\ \bigcap\limits_{n\in\mathbb{N}}(-\infty, x_0-\frac{1}{n}) = (-\infty, x_0)$, поэтому, аналогично п. 2, $F(x_0) = P\{\xi < x_0\} = P\left(\xi^{-1}\left(\bigcap\limits_{n\in\mathbb{N}}(-\infty, x_0 - \frac{1}{n})\right)\right) = \lim\limits_{x\to x_0-0} F(x)$ (здесь вновь показано лишь существование частичного предела, который определяет и существование искомого одностороннего).

4) F монотонна, поэтому $\forall\ x \in \mathbb{R}\ 0 \leq F(x) \leq 1$, значит, функция F может иметь только один скачок, величина которого превышает $\frac{1}{2}$, не более двух скачок, величина которых больше $\frac{1}{3}$, но не превосходит $\frac{1}{2}$; аналогично F имеет не более n скачков, величина которых лежит в пределах $\left[\frac{1}{n+1}, \frac{1}{n}\right)$. Таким образом, всего F имеет конечное или счётное объединение конечного числа скачков, то есть конечное или счётное число скачков. ■

Замечание: некоторые авторы задают функцию распределения как $F_\xi(x) = P\{\xi \leq x\}$; в этом случае свойства 1, 2 и 4 сохраняются, а свойство 3 заменяется на непрерывность справа.

Теорема 3 (без доказательства): любая вещественная функция $F(x)$, удовлетворяющая свойствам 1, 2 и 3 теоремы 2, является функцией распределения некоторой случайной величины (в этом случае $P\{\xi \in [x_1, x_2)\} = F(x_2) - F(x_1)$).

Замечание: таким образом, каждой случайной величине соответствует функция распределения, а каждой функции, удовлетворяющей свойствам 1, 2 и 3 теоремы 2, – случайная величина $\left(\text{возможно, не одна: например, если } \xi = \begin{cases} 1, \ p = \frac{1}{2} \\ -1, \ p = \frac{1}{2}, \end{cases} \text{, то } F_\xi(x) = F_{-\xi}(x)\right)$.

Замечание: функция распределения может служить вероятностной мерой в вероятностном пространстве $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), F_\xi)$, называемом *индуцированным вероятностным пространством.*

Определение: случайная величина ξ называется *распределённой дискретно*, если она принимает конечное или счётное число значений $\{x_1, \dots x_n, \dots\}$; тогда $P\{\xi = x_k\} = p_k > 0$, $\sum\limits_{k \in \mathbb{N}} p_k = 1$. В этом случае часто записывают: $\begin{pmatrix} p_1 \ p_2 \dots p_n \dots \\ x_1 \ x_2 \dots x_n \dots \end{pmatrix}$. Очевидно, что в случае дискретного распределения $F_\xi(x)$ имеет разрывы в точках x_k величиной p_k.

Примеры дискретных распределений:

1) *Вырожденное*: $p\{\xi = c\} = 1$.

2) *Дискретное равномерное*: $\begin{pmatrix} x_1 \dots x_n \\ \frac{1}{n} \dots \frac{1}{n} \end{pmatrix}$.

3) *Бернуллиевское*: $\begin{cases} 0, \ p \\ 1, \ q = 1 - p. \end{cases}$

4) *Биномиальное* с параметрами (n, p): $p_k = C_n^k p^k q^{n-k}$ (вероятность выпадения k успехов в схеме Бернулли, состоящей из n испытаний).

5) *Гипергеометрическое*: $p_m = \dfrac{C_M^m C_{N-M}^{n-m}}{C_N^n}$ (вероятность извлечь m белых шаров в выборке без возвращения n шаров из урны, содержащей N шаров, среди которых M белых).

6) *Геометрическое*: $p_k = pq^k$ (вероятность появления первого успеха в схеме Бернулли после k неудач).

7) *Распределение Паскаля*: $p_k = C_{n+k-1}^k p^n q^k$ (вероятность появления k неудач в схеме Бернулли до n-го успеха).

8) *Распределение Пуассона*: $p_k = \dfrac{\lambda^k e^{-\lambda}}{k!}$, $\lambda > 0$ (приближение схемы Бернулли при $n \to \infty$ в случае $np_n \to \lambda, \ n \to \infty$ – см. теорему Пуассона).

Определение: случайная величина ξ распределена *абсолютно непрерывно*, если $\exists\, p_\xi(x): \ \forall\, x \in \mathbb{R}\, p_\xi(x) \geq 0 : \int\limits_{-\infty}^{\infty} p_\xi(x)dx = 1, \forall\, x \in \mathbb{R}\, F_\xi(x) = \int\limits_{-\infty}^{x} p_\xi(t)dt$. В этом случае $p_\xi(x)$ называется *плотностью распределения случайной величины*; очевидно, $P\{x_1 \leq \xi < x_2\} = \int\limits_{x_1}^{x_2} p_\xi(x)dx, \ \forall\ B \in \mathfrak{B}(\mathbb{R}) \ P\{\xi \in B\} = \int\limits_B p_\xi(x)dx$. Дифференцируя равенство, определяющее $p_\xi(x)$, по x, получим: $\dfrac{dF_\xi}{dx} = p_\xi(x)$; существование $\int\limits_{-\infty}^{x} p_\xi(t)dt \ \forall\ x \in \mathbb{R}$ означает, что $p_\xi = \dfrac{dF_\xi}{dx}$ имеет на $\mathbb{R}$ конечное число точек разрыва, а F_ξ кусочно принадлежит классу $C^1(\mathbb{R})$.

Примеры абсолютно непрерывных распределений:

1) *Равномерное* на $[a,b]$:

$$p_\xi(x)=\begin{cases}\dfrac{1}{b-a},\ x\in[a,b]\\ 0,\ x\notin[a,b]\end{cases}\quad\Rightarrow F_\xi(x)=\begin{cases}0,\ x<a\\ \dfrac{x-a}{b-a},\ a\le x\le b\\ 1,\ x>b\end{cases}$$

(равномерное распределение на $[-\frac{h}{2},\frac{h}{2}]$ описывает ошибку округления при измерении той или иной физической величины, если h – цена деления прибора).

2) *Показательное* с параметром λ:

$$p_\xi(x)=\begin{cases}\lambda e^{-\lambda x},\ x\ge 0\\ 0,\ x<0\end{cases}\quad\Rightarrow F_\xi(x)=\begin{cases}1-e^{-\lambda x},\ x\ge 0\\ 0,\ x<0\end{cases}$$

(описывает время безотказной работы прибора – см. замечание).

3) *Нормальное* с параметрами (a,σ): $p_\xi(x,a,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}}\Rightarrow F_\xi(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{x}e^{-\frac{(u-a)^2}{2\sigma^2}}du$. Нормальное распределение называется *стандартным*, если оно имеет параметры (0,1); в этом случае $F_\xi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}e^{-\frac{u^2}{2}}du=\Phi(x)$. Распределение описывает разброс результатов измерения, вызванный влиянием большого случайных факторов.

Замечание: для показательного распределения характерно *свойство отсутствия последействия*, то есть $\forall\ x,y>0\quad P\{\xi\ge x+y/\xi\ge x\}=\dfrac{P\{\xi\ge x+y,\ \xi\ge x\}}{P\{\xi\ge x\}}=$ $=\dfrac{1-P\{\xi<x+y\}}{1-P\{\xi<x\}}=\dfrac{e^{-\lambda(x+y)}}{e^{-\lambda x}}=e^{-\lambda y}=P\{\xi\ge y\}$. Это означает, что вероятность работы прибора в течение некоторого времени после того, как он уже проработал время t_0, не зависит от величины t_0.

Определение: пусть есть $n\ge 2$ случайных величин с функциями распределения $F_1(x),\dots F_n(x)$. $F(x)$ называется *смесью распределений*, задаваемых $F_1,\dots F_n$, если $F(x)=$ $=\sum\limits_{k=1}^{n}\alpha_k F_k(x)$ $(\alpha_k\ge 0\ \forall\ k=\overline{1,n},\ \sum\limits_{k=1}^{n}\alpha_k=1)$.

Теорема 4 (Лебега, без доказательства): любая функция распределения $F(x)$ представима в виде смеси распределений, задаваемых $F_1(x),\ F_2(x),\ F_3(x)$, где F_1 – функция распределения дискретно распределённой случайной величины, F_2 – функция распределения случайной величины, распределённой абсолютно непрерывно, F_3 – *сингулярная функция распределения* (та, для которой нельзя задать плотность распределения).

3.2. Многомерные случайные величины.

Определение: $(\Omega,\mathfrak{F},P)$ – вероятностное пространство; $\xi_1,\dots\xi_n$: $\Omega\to\mathbb{R}$ – случайные величины $(n\ge 2)$; тогда $(\xi_1,\dots\xi_n)$: $\Omega^n\to\mathbb{R}$ называется *случайным вектором.* $F_{\xi_1\dots\xi_n}(x_1,\dots x_n)=P\{\xi_1<x_1,\dots\xi_n<x_n\}$ – *многомерная совместная функция распределения* $\xi_1,\dots\xi_n$ (функция распределения случайного вектора).

Теорема 1 (свойства многомерной функции распределения): $\overrightarrow{\xi}=(\xi_1,\dots\xi_n)$ – случайный вектор; $F_{\xi_1,\dots\xi_n}(x_1,\dots x_n)$ – его функция распределения; тогда

1) $\forall\ x_{11},\ x_{12},\ x_2,\dots,x_n\in\mathbb{R}:x_{11}<x_{12}\ F(x_{11},x_2,\dots x_n)\le F(x_{12},x_2,\dots x_n)$.

2) $\forall\ k=\overline{1,n}\ \lim\limits_{x_k\to-\infty}F(x_1,\dots x_n)=0,\ \lim\limits_{x_1,\dots x_n\to+\infty}F(x_1,\dots x_n)=1.$

3) $\forall\ x_1,\dots x_n\in\mathbb{R},\ \forall\ k=\overline{1,n}\ \lim\limits_{x_k\to x_{k_0}-0}F(x_1,\dots x_n)=F(x_1,\dots x_{k-1},x_{k_0},x_{k+1},\dots x_n).$

4) $\forall\ a_i,b_i\in\mathbb{R}:a_i<b_i(i=\overline{1,n})\ P\{a_1\le\xi_1<b_1,\dots a_n\le\xi_n<b_n\}=F(b_1,\dots b_n)-\sum\limits_{i=1}^{n}p_i+\sum\limits_{i<j}p_{ij}-\ldots+(-1)^nF(a_1,\dots a_n)\ge0$, где p_{ij} – значение F в точке, для которой $i,j,\dots$-ые координаты берутся как $a_i,a_j,\dots$, а остальные как $b_l(l\neq i,j,\dots)$. В частности, при $n=2$ $P\{a_1\le b_1,a_2\le\xi_2<b_2\}=F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)+F(a_1,a_2)$.

△ Свойства 1–3 следуют из соответствующих свойств для одномерной функции распределения (см. 3.1, теорема 2), а свойство 4 оставим без доказательства. ■

Определение: случайный вектор $\overrightarrow{\xi}$ распределён *дискретно*, если функция $\xi\colon\Omega\to\mathbb{R}^n$ принимает конечное или счётное число значений $x_s=(x_{s1},\dots x_{sn})$; тогда $P\{\xi=x_s\}=p_s\ge0\ \forall\ s,\ \sum\limits_s p_s=1$. $\overrightarrow{\xi}$ распределён *абсолютно непрерывно*, если $\exists\ p_{\xi_1\dots\xi_n}(x_1,\dots x_n)\colon\mathbb{R}^n\to\mathbb{R}$: $\forall\ (x_1,\dots x_n)\in\mathbb{R}^n\ p(x_1,\dots x_n)\ge0,\int\limits_{\mathbb{R}^n}p(x_1,\dots x_n)dx_1\dots dx_n=1$

$$F_{\xi_1\dots\xi_n}=\int\limits_{-\infty}^{x_1}\cdots\int\limits_{-\infty}^{x_n}p(u_1,\dots u_n)du_1\dots du_n;$$

таким образом, $p(x_1,\dots x_n)=\dfrac{\partial^nF}{\partial x_1\dots\partial x_n}$.

Определение: если $\overrightarrow{\xi}=(\xi_1,\xi_2)$ – случайный вектор, то $F_{\xi_1}(x),\ F_{\xi_2}(y)$ называются *маргинальными* (одномерным) функциями распределения; аналогично, $p_{\xi_1},\ p_{\xi_2}$ – *маргинальные плотности распределения.* Очевидно, что

$$F_{\xi_1}(x)=\int\limits_{-\infty}^{x}\left(\int\limits_{-\infty}^{+\infty}p_{\xi_1\xi_2}(u,v)dv\right)du\Rightarrow\ p_{\xi_1}(x_1)=\int\limits_{\mathbb{R}}p_{\xi_1\xi_2}(x,v)dv.$$

Аналогично в n-мерном случае маргинальные плотности получаются интегрированием по $n-1$ переменным.

Примеры многомерных распределений:

1) *Равномерное двумерное* на области D: $p_{\xi_1\xi_2}(x,y)=\begin{cases}0,\ (x,y)\notin D\\ \frac{1}{S(D)},\ (x,y)\in D.\end{cases}$

2) *Двумерное нормальное* с параметрами $(a_1,a_2,\sigma_1,\sigma_2,\rho)$:

$$p_{\xi_1\xi_2}(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{(x-a_1)^2}{2\sigma_1^2(1-\rho^2)}+\frac{2\rho(x-a_1)(y-a_2)}{\sigma_1\sigma_2(1-\rho^2)}-\frac{(y-a_2)^2}{2\sigma_2^2(1-\rho^2)}\right).$$

Можно показать, что в этом случае ρ имеет смысл коэффициента корреляции случайных величин ξ_1 и ξ_2 (см. 3.4), а

$$p_{\xi_1}(x)=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(x-a_1)^2}{2\sigma_1^2}\right),\ p_{\xi_2}(y)=\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(y-a_2)^2}{2\sigma_2^2}\right).$$

Определение: случайные величины $\xi_1,\dots\xi_n$ *независимы*, если

1) $F_{\xi_1\dots\xi_n}(x_1,\dots x_n)=F_{\xi_1}(x_1)\cdot\ldots\cdot F_{\xi_n}(x_n)$ во всех точках $(x_1,\dots x_n)$, в которых эти функции определены.

2) $\forall\ B_1, \dots B_n \in \mathfrak{B}(\mathbb{R})\ P\{\xi_1 \in B_1, \dots \xi_n \in B_n\} = P\{\xi_1 \in B_1\} \cdot \ldots \cdot P\{\xi_n \in B_n\}$.

Замечание: по теореме 1 (3.1) данные определения эквиваленты; отметим также, что на независимость случайных величин в совокупности накладывается меньшее число условий, чем на независимость в совокупности событий. Это связано с тем, что совместная функция распределения n случайных величин позволяет задать совместные функции распределения меньшего числа случайных величин, а также маргинальные функции распределения.

Теорема 2 (без доказательства): случайный вектор $(\xi_1, \dots \xi_n)$ распределён дискретно; тогда для того, чтобы $\xi_1, \dots \xi_n$ были независимы, необходимо и достаточно $P\{\xi_1 = x_1, \dots \xi_n = x_n\} = P\{\xi_1 = x_1\} \cdot \ldots \cdot P\{\xi_n = x_n\}\ \forall\ (x_1, \dots x_n) \in \mathbb{R}^n$.

Следствие: случайные величины, одна из которых имеет вырожденное распределение, независимы (равенство, заданное в условии теоремы, выполняется во всех точках $\mathbb{R}^2$).

Теорема 3 (без доказательства): случайный вектор $(\xi_1, \dots \xi_n)$ распределён абсолютно непрерывно; тогда для того, чтобы $\xi_1, \dots \xi_n$ были независимы, необходимо и достаточно $p_{\xi_1 \dots \xi_n}(x_1, \dots x_n) = p_{\xi_1}(x_1) \cdot \ldots \cdot p_{\xi_n}(x_n)$ во всех точках непрерывности этих плотностей.

3.3. Функции от случайных величин.

Определение: g: $\mathbb{R}^n \to \mathbb{R}^m$ – *борелевская функция*, если $\forall\ B \in \mathfrak{B}(\mathbb{R}^m)\ g^{-1}(B) \in \mathfrak{B}(\mathbb{R}^n)$. Таким образом, если ξ – случайный вектор, то $\eta = g(\xi)$ – также случайный вектор.

Распределение функции от случайной величины: пусть $\overrightarrow{\xi}$ – случайный вектор, g: $\mathbb{R}^n \to \mathbb{R}^n$ – борелевская функция, $\overrightarrow{\eta} = g(\overrightarrow{\xi})$; тогда $\forall\ B \in \mathfrak{B}(\mathbb{R}^n)$

$$P\{\eta \in B\} = P\{\xi \in g^{-1}(B)\} = \int\limits_{g^{-1}(B)} p_\xi(x)dx = (\text{замена переменных в кратном интеграле}) =$$

$$= \int\limits_B p_\xi(g^{-1}(y))|J(g^{-1}(y))|^{-1}dy, \text{ где } J = \det\left\{\frac{\partial g_i}{\partial x_j}\right\}_{i,j=1}^n.$$

Пример (распределение χ^2 с одной степенью свободы): случайная величина ξ распределена нормально с парметрами (0,1); тогда $\eta = \xi^2$ имеет *распределение хи-квадрат* с одной степенью свободы. Найдём плотность распределения η.

$$F_\eta(y) = P\{\eta < y\} = \begin{cases} P\{\xi^2 < y\},\ y > 0 \\ 0,\ y \le 0 \end{cases} = \begin{cases} P\{-\sqrt{y} < \xi < \sqrt{y}\},\ y > 0 \\ 0,\ y \le 0 \end{cases} =$$

$$= \begin{cases} \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\sqrt{y}}^{\sqrt{y}} e^{-\frac{x^2}{2}}dx,\ y > 0 \\ 0,\ y \le 0 \end{cases} = \begin{cases} \sqrt{\dfrac{2}{\pi}}\Phi_0(\sqrt{y}),\ y > 0 \\ 0,\ y \le 0, \end{cases}$$

$$\text{где } \Phi_0(y) = \int\limits_0^y e^{-\frac{x^2}{2}}dx \Rightarrow p_\eta(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}},\ y > 0 \\ 0,\ y \le 0. \end{cases}$$

Теорема 1 (без доказательства): пусть случайные величины $\xi_1^{(i)}, \dots \xi_{n_i}^{(i)}$ $(i = \overline{1,m})$ – независимые случайные величины; φ_i: $\mathbb{R}^{n_i} \to R$ – борелевские функции; тогда случайные величины $\eta_i = \varphi_i(\xi_1^{(i)}, \dots \xi_{n_i}^{(i)})$ также независимы.

Теорема 2 (формула композиции (свёртки)): ξ, η – независимые случайные величины, распределённые абсолютно непрерывно; $\tau = \xi + \eta \Rightarrow p_\tau(z) = \int\limits_{\mathbb{R}} p_\xi(u) p_\eta(z-u)du$.

$$\triangle\ F_\tau(z) = F_{\xi+\eta}(z) = \iint\limits_{u+v<z} p_{\xi\eta}(u,v)dudv = \int\limits_{-\infty}^{z-u} dv \int\limits_{-\infty}^{+\infty} p_{\xi\eta}(u,v)du = \begin{pmatrix} u = u \\ v = w-u \end{pmatrix} =$$

$$= \int\limits_{-\infty}^{z} \left(\int\limits_{-\infty}^{+\infty} p_{\xi\eta}(u, w-u)du \right) dw = \int\limits_{-\infty}^{z} p_\tau(w)dw \Rightarrow p_\tau(z) = \int\limits_{-\infty}^{+\infty} p_{\xi\eta}(u, z-u)du =$$

$$= (\xi,\ \eta \text{ независимы}) = \int\limits_{-\infty}^{+\infty} p_\xi(u)p_\eta(z-u)du. \ \blacksquare$$

Следствие: $\xi,\ \eta$ распределены нормально с параметрами (a_1, σ_1), (a_2, σ_2) соответственно; тогда случайная величина $\tau = \xi + \eta$ распределена нормально с параметрами $(a_1 + a_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

$$\triangle \text{ По формуле композиции } p_\tau(z) = \frac{1}{2\pi\sigma_1\sigma_2} \int\limits_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(\frac{(x-a_1)^2}{\sigma_1^2} + \frac{(z-x-a_2)^2}{\sigma_2^2}\right)\right) dx =$$

$$= \begin{pmatrix} u = z - a_1 - a_2 \\ v = x - a_1 \end{pmatrix} = \frac{1}{2\pi\sigma_1\sigma_2} \int\limits_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(\frac{v^2}{\sigma_1^2} + \frac{(u-v)^2}{\sigma_2^2}\right)\right) dv =$$

$$= \frac{1}{2\pi\sigma_1\sigma_2} \int\limits_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(v^2 \cdot \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2} - 2uv \cdot \frac{1}{\sigma_2^2} + \frac{u^2}{\sigma_2^2} \cdot \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2} - \frac{u^2}{\sigma_2^2} \cdot \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2} + \frac{u^2}{\sigma_2^2}\right)\right) dv =$$

$$= \frac{1}{2\pi\sigma_1\sigma_2} \int\limits_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(\left(\frac{u\sigma_1}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}} - v \cdot \frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}\right)^2 + \frac{u^2}{\sigma_2^2}\left(1 - \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)\right)\right) dv =$$

$$= \left(t = v \cdot \frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2} - \frac{u\sigma_1}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}},\ |J| = \frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}\right) = \frac{1}{2\pi\sigma_1\sigma_2} \cdot e^{-\frac{u^2}{2(\sigma_1^2+\sigma_2^2)}}.$$

$$\cdot \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} \sqrt{2\pi} \cdot \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\ e^{-\frac{t^2}{2}} dt = \left(\frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} e^{-\frac{t^2}{2}} dt = 1\right) = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \cdot e^{-\frac{(z-a_1-a_2)^2}{2(\sigma_1^2+\sigma_2^2)}}$$

– нормальное распределение с параметрами $(a_1 + a_2, \sqrt{\sigma_1^2+\sigma_2^2})$. $\blacksquare$

Пример: $\xi,\ \eta$ – независимые случайные величины; ξ распределена нормально с параметрами $(0, \sigma)$, а η – равномерно на отрезке $[-\frac{h}{2}, \frac{h}{2}]$. Найдём $P_\tau(x)$, где $\tau = \xi + \eta$.

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}},\ p_\eta(x) = \begin{cases} \frac{1}{h},\ x \in [-\frac{h}{2}, \frac{h}{2}] \\ 0,\ x \notin [-\frac{h}{2}, \frac{h}{2}] \end{cases} \Rightarrow$$

$$\Rightarrow p_\tau(x) = \int\limits_{x-\frac{h}{2}}^{x+\frac{h}{2}} \frac{1}{h\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt = \frac{1}{h}\left(\Phi_0\left(\frac{x+\frac{h}{2}}{\sigma}\right) - \Phi_0\left(\frac{x-\frac{h}{2}}{\sigma}\right)\right).$$

Замечание: независимые случайные величины могут быть связаны функционально; например, в случае вырожденного распределения ξ в точке $x = 1$, $\eta = \xi^2$ имеет такое же распределение, а константы независимы (см. 3.2, следствие из теоремы 2).

Определение: $\overrightarrow{\xi} = (\xi_1, \dots \xi_n)$, $\overrightarrow{\eta} = (\eta_1, \dots \eta_m)$ – случайные вектора; $D_\xi \subset \mathbb{R}^n$, $P\{\xi \in D_\xi\} > 0$; тогда

$$F(y_1, \dots y_m) = P\{\eta_1 < y_1, \dots \eta_m < y_m | \xi \in D_\xi\} = \frac{P\{\eta_1 < y_1, \dots \eta_m < y_m, \xi \in D_\xi\}}{P\{\xi \in D_\xi\}}$$

– *условная функция распределения* случайной величины η.

Если ξ и η распределены *дискретно*, то можно обозначить $P\{\xi = x_k, \eta = y_m\} = p_{km}$; тогда $P\{\xi = x_k\} = \sum\limits_m p_{km} = p_k \Rightarrow P\{\eta = y_m | \xi = x_k\} = \dfrac{p_{km}}{p_k}$ – вероятность для η при фиксированном ξ.

Если ξ и η распределены *абсолютно непрерывно*, то

$$P\{\eta < y | x \leq \xi < x + h\} = \frac{\int\limits_{-\infty}^{y} \int\limits_{x}^{x+h} p_{\xi\eta}(u, v) du dv}{\int\limits_{x}^{x+h} \left(\int\limits_{-\infty}^{+\infty} p_{\xi\eta}(u, v) dv \right) du} \xrightarrow[h \to 0]{} \frac{\int\limits_{-\infty}^{y} p_{\xi\eta}(x, v) dv}{p_\xi(x)} = P\{\eta < y | \xi = x\}$$

– *функция распределения η при фиксированном ξ.*

3.4. Числовые характеристики случайных величин.

Определение: ξ – случайная величина, распределённая дискретно $\begin{pmatrix} x_1 \dots x_n \dots \\ p_1 \dots p_n \dots \end{pmatrix}$, тогда $\mathbf{M}\xi = \sum\limits_k x_k p_k$ – *математическое ожидание* ξ (существует, если соответствующий ряд сходится абсолютно); если ξ – случайная величина, распределённая абсолютно непрерывно, то $\mathbf{M}\xi = \int\limits_{-\infty}^{+\infty} x p_\xi(x) dx$ (существует, если соответстующий несобственный интеграл сходится абсолютно). Если же ряд (интеграл) сходится условно или расходится, то математическое ожидание ξ не существует.

Определение: $\overrightarrow{\xi} = (\xi_1, \dots \xi_n)$ – случайный вектор; g: $\mathbb{R}^n \to \mathbb{R}$ – борелевская функция; $\eta = g(\xi)$. Если ξ распределена дискретно по векторам $x_1, \dots x_n, \dots$, то $\mathbf{M}\eta \overset{def}{=} \sum\limits_k g(x_k) p_k$; если ξ распределена абсолютно непрерывно, то $\mathbf{M}\eta \overset{def}{=} \int\limits_{\mathbb{R}^n} g(x_1, \dots x_n) p_{\xi_1 \dots \xi_n}(x_1, \dots x_n) \cdot$ $\cdot dx_1 \dots dx_n$ (для существования математического ожидания ξ ряд (интеграл) должен сходиться абсолютно).

Теорема 1 (свойства математического ождиания): ξ, η – случайные величины, имеющие математическое ождиание; тогда

1) $\forall\ C \in \mathbb{R}$ $\mathbf{M}(C\xi) = C \cdot \mathbf{M}\xi$;
2) $\forall\ C \in \mathbb{R}$ $\mathbf{M}C = C$;
3) $\mathbf{M}\xi \leq \mathbf{M}|\xi|$;
4) $\mathbf{M}(\xi + \eta) = \mathbf{M}\xi + \mathbf{M}\eta$;
5) Если ξ и η независимы, то $\mathbf{M}(\xi\eta) = \mathbf{M}\xi \cdot \mathbf{M}\eta$.

$\triangle$ Доказательства свойств 1-3 следуют из соответствующих свойств рядов и несобственных интегралов.

$$4)\ \mathbf{M}(\xi + \eta) = \int\limits_{\mathbb{R}^2} (x + y) p_{\xi\eta}(x, y) dx dy = \int\limits_{\mathbb{R}^2} x p_{\xi\eta}(x, y) dx dy + \int\limits_{\mathbb{R}^2} y p_{\xi\eta}(x, y) dx dy =$$

$$= \int\limits_{\mathbb{R}} x \left(\int\limits_{\mathbb{R}} p_{\xi\eta}(x,y)dy \right) dx + \int\limits_{\mathbb{R}} y \left(\int\limits_{\mathbb{R}} p_{\xi\eta}(x,y)dx \right) dy =$$

$$= \int\limits_{\mathbb{R}} xp_{\xi}(x)dx + \int\limits_{\mathbb{R}} yp_{\eta}(y)dy = \mathbf{M}\xi + \mathbf{M}\eta;$$

$$5)\ \mathbf{M}(\xi\eta) = \int\limits_{\mathbb{R}^2} xyp_{\xi\eta}(x,y)dxdy = \int\limits_{\mathbb{R}} xp_{\xi}(x)dx \cdot \int\limits_{\mathbb{R}} yp_{\eta}(y)dy = \mathbf{M}\xi \cdot \mathbf{M}\eta.$$

$$(p_{\xi\eta}(x,y) = p_{\xi}(x) \cdot p_{\eta}(y)) . \blacksquare$$

Замечание: если случайная величина η может быть представлена в виде $\eta = \xi_1 + \ldots + \xi_n$ (где ξ_i независимы, одинаково распределены и называются *индикаторами*), то $\mathbf{M}\eta = \sum\limits_{i=1}^{n} \mathbf{M}\xi_i$.

Примеры:

1) *Биномиальное распределение*: μ_n распределена биномиально с параметрами (n, p); тогда $\mu_n = \xi_1 + \ldots + \xi_n$, где ξ_i имеют распределение Бернулли с параметром p; тогда из замечания следует, что $\mathbf{M}\mu_n = np$.

2) *Геометрическое распределение*:

$$p_k = pq^k \Rightarrow \mathbf{M}\xi = \sum_{k=1}^{\infty} kpq^k = pq \sum_{k=1}^{\infty} kq^{k-1} = pq \left(\sum_{k=1}^{\infty} q^k \right)' = pq \left(\frac{q}{1-q} \right)' = \frac{q}{p}.$$

3) *Распределение Паскаля*: $\mu_n = \xi_1 + \ldots + \xi_{n+k}$, где индикаторы ξ_i имеют геометрическое распределение; тогда из замечания следует, что $\mathbf{M}\mu_n = \sum\limits_{i=1}^{n+k} \mathbf{M}\xi_i = \dfrac{q}{p}(n+k)$.

4) *Гипергеометрическое распределение*: $\mu_n = \xi_1 + \ldots + \xi_n$, где ξ_i принимает значение 1 при вытаскивании белого шара и значение 0 при вытаскивании чёрного. $\mathbf{M}\xi_i = \dfrac{M}{N}$; тогда $\mathbf{M}\mu_n = \dfrac{nM}{N}$.

5) *Распределение Пуассона*: $p_k = \dfrac{\lambda^k e^{-\lambda}}{k!} \Rightarrow \mathbf{M}\xi = \lambda e^{-\lambda} \cdot \sum\limits_{k=1}^{+\infty} \dfrac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda$.

6) *Равномерное расределение*: $\mathbf{M}\xi = \int\limits_a^b \dfrac{xdx}{b-a} = \dfrac{(b^2-a^2)}{2(b-a)} = \dfrac{a+b}{2}$.

7) *Нормальное распределение*:

$$\mathbf{M}\xi = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{\mathbb{R}} x \exp\left(-\frac{(x-a)^2}{2\sigma^2} \right) dx = \frac{1}{\sqrt{2\pi}\sigma} \cdot \frac{1}{2} \int\limits_{\mathbb{R}} \exp\left(-\frac{(x-a)^2}{2\sigma^2} \right) d(x-a)^2 +$$

$$+ \frac{a}{\sqrt{2\pi}\sigma} \cdot \int\limits_{\mathbb{R}} \exp\left(-\frac{(x-a)^2}{2\sigma^2} \right) dx = 0 + a \cdot 1 = a.$$

8) *Показательное распределение*: $\mathbf{M}\xi = \int\limits_0^{+\infty} x\alpha e^{-\alpha x} dx = -xe^{-\alpha x}|_0^{+\infty} + \int\limits_0^{+\infty} e^{-\alpha x} dx = \dfrac{1}{\alpha}$.

9) *Распределение Коши*: $p_{\xi}(x) = \dfrac{1}{\pi} \dfrac{1}{1+x^2} \Rightarrow \dfrac{1}{\pi} \int\limits_{\mathbb{R}} \dfrac{|x|dx}{1+x^2}$ – расходится, то есть математическое ожидание распределения Коши не существует.

Определение: ξ – случайная величина, распределённая дискретно и имеющая математическое ожидание; тогда $\mathbf{D}\xi = \mathbf{M}\left((\xi - \mathbf{M}\xi)^2\right) = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2$ называется *дисперсией* ξ, а $\sqrt{\mathbf{D}\xi}$ – *средним квадратическим отклонением* ξ (определены только в том случае, когда существует $\mathbf{M}(\xi - \mathbf{M}\xi)^2$. Очевидно, что если ξ распределена дискретно, то $\mathbf{D}\xi = \sum\limits_k (x_k - \mathbf{M}\xi)^2 p_k$, а если ξ распределена абсолютно непрерывно, то $\mathbf{D}\xi = \int\limits_{\mathbb{R}} (x - \mathbf{M}\xi)^2 p_\xi(x)dx = \int\limits_{\mathbb{R}} x^2 p_\xi(x)dx - (\mathbf{M}\xi)^2$.

Теорема 2 (свойства дисперсии): ξ, η – случайные величины, имеющие дисперсию; тогда

1) $\mathbf{D}\xi \geq 0$;
2) $\forall\ C \in \mathbb{R}\ \mathbf{D}C = 0$;
3) $\forall\ C \in \mathbb{R}\ \mathbf{D}(C\xi) = C^2 \cdot \mathbf{D}\xi$;
4) $\mathbf{D}(\xi_1 + \xi_2) = \mathbf{D}\xi_1 + \mathbf{D}\xi_2 + 2\mathbf{M}(\xi_1 - \mathbf{M}\xi_1)(\xi_2 - \mathbf{M}\xi_2)$; если ξ_1 и ξ_2 независимы, то $\mathbf{D}(\xi_1 + \xi_2) = \mathbf{D}\xi_1 + \mathbf{D}\xi_2$.

$\triangle$ Доказательства свойств 1-3 следуют из определения дисперсии, а также свойств числовых рядов и несобственных интегралов.

4) $\mathbf{D}(\xi_1 + \xi_2) = \mathbf{M}(\xi_1 + \xi_2 - \mathbf{M}(\xi_1 + \xi_2))^2 = \mathbf{M}\left((\xi_1 - \mathbf{M}\xi_1) + (\xi_2 - \mathbf{M}\xi_2)\right)^2 = \mathbf{D}\xi_1 + \mathbf{D}\xi_2 + 2\mathbf{M}(\xi_1 - \mathbf{M}\xi_1)(\xi_2 - \mathbf{M}\xi_2)$. Если ξ_1 и ξ_2 независимы, то $\mathbf{M}\left((\xi_1 - \mathbf{M}\xi_1)(\xi_2 - \mathbf{M}\xi_2)\right) = (\mathbf{M}\xi_1 - \mathbf{M}\xi_1)(\mathbf{M}\xi_2 - \mathbf{M}\xi_2) = 0 \Rightarrow \mathbf{D}(\xi_1 + \xi_2) = \mathbf{D}\xi_1 + \mathbf{D}\xi_2$. $\blacksquare$

Примеры:

1) *Биномиальное распределение*: $\mu_n = \xi_1 + \ldots + \xi_n$, ξ_i имеют распределение Бернулли с параметром p; тогда $\mathbf{D}\xi_i = \mathbf{M}\xi_i^2 - (\mathbf{M}\xi_i)^2 = p - p^2 \Rightarrow$ (ξ_i независимы) $\mathbf{D}\mu_n = n(p - p^2) = npq$.

2) *Распределение Пуассона*: $\mathbf{M}\xi^2 - \mathbf{M}\xi = \mathbf{M}(\xi(\xi - 1)) = \sum\limits_{m=0}^{+\infty} m(m-1) \cdot \dfrac{\lambda^m}{m!} e^{-\lambda} = \lambda^2 \cdot \sum\limits_{m=2}^{+\infty} \dfrac{\lambda^{m-2}}{(m-2)!} \cdot e^{-\lambda} = \lambda^2 \Rightarrow \mathbf{M}\xi^2 = \lambda^2 + \lambda \Rightarrow \mathbf{D}\xi = \lambda$ (под $-2!$ понимается 2, под $-1!$ (-1)).

3) *Гипергеометрическое распределение*: $\mu_n = \xi_1 + \ldots + \xi_n$, ξ_i – случайная величина, принимающая значение 1 при извлечении белого шара и 0 при извлечении чёрного, то есть

$$\xi_i = \begin{cases} 1, & \frac{M}{N} \\ 0, & 1 - \frac{M}{N} \end{cases}; \mathbf{M}\mu_n = n\frac{M}{N};\ \mathbf{M}\mu_n^2 = \mathbf{M}\left(\sum_{k=1}^{n} \xi_k\right)^2 = \sum_{k=1}^{n} \mathbf{M}\xi_k^2 + \sum_{k \neq l} \mathbf{M}\xi_k\xi_l = n\frac{M}{N} +$$

$$+ n(n-1)\frac{M(M-1)}{N(N-1)};\ \mathbf{D}\mu_n = \mathbf{M}\mu_n^2 - (\mathbf{M}\mu_n)^2 = n\frac{M}{N} + n(n-1)\frac{M(M-1)}{N(N-1)} - n^2\frac{M^2}{N^2} =$$

$$= n\frac{M}{N} \cdot \frac{N(N-1) + MN(M-1)(n-1) - nM^2(N-1)}{N(N-1)} = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

4) *Равномерное распределение*:

$$\mathbf{D}\xi = \int_a^b \frac{(x - \frac{a+b}{2})^2}{b-a} dx = \frac{1}{3(b-a)}\left(x - \frac{a+b}{2}\right)^3 \Big|_a^b = \frac{1}{3(b-a)}\left(\frac{(b-a)^3}{8} - \frac{(a-b)^3}{8}\right) = \frac{(b-a)^2}{12}.$$

5) *Нормальное распределение*: $\mathbf{D}\xi = \dfrac{1}{\sqrt{2\pi}\sigma} \int\limits_{\mathbb{R}} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left(y = \dfrac{x-a}{\sigma}\right) = \dfrac{\sigma^2}{\sqrt{2\pi}} \cdot \int\limits_{\mathbb{R}} y^2 e^{-\frac{y^2}{2}} dy = -\dfrac{\sigma^2}{\sqrt{2\pi}}\left(y e^{-\frac{y^2}{2}} \Big|_{-\infty}^{+\infty} - \int\limits_{\mathbb{R}} e^{-\frac{y^2}{2}} dy\right) = \dfrac{\sigma^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = \sigma^2$.

Замечание: ξ – случайная величина, имеющая мат. ожидание и дисперсию; $\eta = \dfrac{\xi - \mathbf{M}\xi}{\sqrt{\mathbf{D}\xi}}$ $\Rightarrow \mathbf{M}\eta = \dfrac{\mathbf{M}(\xi - \mathbf{M}\xi)}{\sqrt{\mathbf{D}\xi}} = \dfrac{\mathbf{M}\xi - \mathbf{M}\xi}{\sqrt{\mathbf{D}\xi}} = 0,\ \mathbf{D}\eta = \dfrac{\mathbf{D}(\xi - \mathbf{M}\xi)}{\mathbf{D}\xi} = \dfrac{\mathbf{D}\xi}{\mathbf{D}\xi} = 1$; таким образом получим η – *центрированную и нормированную* случайную величину.

Определение: ξ и η – случайные величины; тогда $\operatorname{cov}(\xi,\eta) \stackrel{def}{=} \mathbf{M}\left((\xi - \mathbf{M}\xi)(\eta - \mathbf{M}\eta)\right) = \mathbf{M}(\xi\eta) - \mathbf{M}\xi \cdot \mathbf{M}\eta$ – *ковариация* случайных величин ξ и η.

Теорема 3 (свойства ковариации): $\xi,\ \eta$ – случайные величины; тогда

1) $\operatorname{cov}(\xi,\xi) = \mathbf{D}\xi$;

2) $\operatorname{cov}(\xi,\eta) = \operatorname{cov}(\eta,\xi)$;

3) $\forall\ C_1,\ C_2 \in \mathbb{R}\ \operatorname{cov}(C_1\xi, C_2\eta) = C_1C_2\operatorname{cov}(\xi,\eta)$;

4) ξ и η независимы; тогда $\operatorname{cov}(\xi,\eta) = 0$ (обратное неверно).

$\triangle$ Все свойства следуют из определения ковариации, дисперсии, математического ожидания и свойств мат. ожидания.

Теорема 4: $\xi_1, \dots \xi_n$ – случайные величины; тогда $\forall\ c_1, \dots c_n \in \mathbb{R}$

$$\mathbf{D}\left(\sum_{i=1}^{n} c_i\xi_i\right) = \sum_{i,j=1}^{n} c_ic_j\operatorname{cov}(\xi_i,\xi_j).$$

$\triangle$ Покажем, что равенство верно при $n = 2$:

$$\mathbf{D}(c_1\xi_1 + c_2\xi_2) = \mathbf{M}\left((c_1\xi_1 - c_1\mathbf{M}\xi_1)^2 + (c_2\xi_2 - c_2\mathbf{M}\xi_2)^2 + 2(c_1\xi_1 - c_1\mathbf{M}\xi_1)(c_2\xi_2 - c_2\mathbf{M}\xi_2)\right) = \\ = c_1^2\mathbf{D}\xi_1 + c_2^2\mathbf{D}\xi_2 + 2c_1c_2\operatorname{cov}(\xi_1,\xi_2);$$

аналогично можно по индукции перейти к общему утверждению. $\blacksquare$

Определение: $\xi,\ \eta$ – случайные величины; тогда $\rho(\xi,\eta) \stackrel{def}{=} \dfrac{\operatorname{cov}(\xi,\eta)}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}}$ – *коэффициент корреляции* величин ξ и η.

Теорема 5 (свойства коэффициента корреляции): $\xi,\ \eta$ – случайные величины; тогда

1) $|\rho(\xi,\eta)| \leq 1$;

2) Если ξ и η независимы, то $\rho(\xi,\eta) = 0$;

3) $\forall\ c_1, c_2 \in \mathbb{R},\ \eta = c_1\xi + c_2\ |\rho(\xi,\eta)| = 1$.

$\triangle$ 1) $\forall\ \lambda \in \mathbb{R}\ \mathbf{D}(\lambda\xi + \eta) = \lambda^2\mathbf{D}\xi + 2\lambda\operatorname{cov}(\xi,\eta) + \mathbf{D}\eta \geq 0 \Rightarrow$ (рассматриваем как квадратный трёхчлен относительно λ) $\Rightarrow (\operatorname{cov}(\xi,\eta))^2 - \mathbf{D}\xi \cdot \mathbf{D}\eta \leq 0 \Rightarrow |\operatorname{cov}(\xi,\eta)| \leq \sqrt{\mathbf{D}\xi \cdot \mathbf{D}\eta} \Rightarrow |\rho(\xi,\eta)| \leq 1$.

2) Следует из свойства 4 теоремы 3.

3) Пусть $\mathbf{M}\xi = a,\ \mathbf{D}\xi = \sigma^2 \Rightarrow \mathbf{M}\eta = c_1a + c_2,\ \mathbf{D}\eta = c_1^2\sigma^2,\ \operatorname{cov}(\xi,\eta) = \mathbf{M}(\xi - a)(c_1\xi + c_2 - c_1a - c_2) = c_1\mathbf{M}(\xi - a)^2 = c_1\sigma^2 \Rightarrow |\rho(\xi,\eta)| = \left|\dfrac{c_1\sigma^2}{c_1\sigma^2}\right| = 1$. $\blacksquare$

Определение: случайные величины ξ и η называются некоррелированными, если $\rho(\xi,\eta) = 0$, и коррелированными в обратном случае.

Замечание: двумерное нормальное распределение случайных величин ξ,η невырождено в случае $|\rho(\xi,\eta)| < 1$ и вырождено при $|\rho(\xi,\eta)| = 1$; при $\rho = 0$ случайные величины ξ_1 и ξ_2 двумерного нормального распределения независимы.

Определение: *смешанным моментом* порядка k называется $\mathbf{M}\xi^k$, *абсолютным моментом* порядка k: $\mathbf{M}|\xi|^k$. *Центральный момент* порядка k: $\mathbf{M}(\xi - \mathbf{M}\xi)^k$, *абсолютный центральный момент* порядка k: $\mathbf{M}|\xi - \mathbf{M}\xi|^k$.

Замечание: если существует момент k-го порядка случайной величины ξ, то существуют и все моменты более низких порядков этой случайной величины.

3.5. Последовательности случайных величин.

Теорема 1 (неравенство Маркова): $\xi \geq 0$ – случайная величина, имеющая математическое ожидание; $a \geq 0 \Rightarrow P\{\xi \geq a\} \leq \dfrac{\mathbf{M}\xi}{a}$.

$\triangle$ Докажем теорему для двух случаев – абсолютно непрерывного и дискретного распределений ξ.

1) Абсолютно непрерывное распределение ξ: $p_\xi(x) = \begin{cases} 0, \ x < 0 \\ \varphi(x), \ x \geq 0, \end{cases}$ $\int\limits_0^{+\infty} \varphi(x)dx = 1 \Rightarrow$

$\forall\ a > 0\ \mathbf{M}\xi = \int\limits_0^{+\infty} x\varphi(x)dx \geq \int\limits_a^{+\infty} x\varphi(x)dx \geq a \cdot \int\limits_a^{+\infty} \varphi(x)dx = a \cdot P\{\xi \geq a\} \Rightarrow P\{\xi \geq a\} \leq \dfrac{\mathbf{M}\xi}{a}$.

2) Дискретное распределение ξ: пусть ξ принимает значения $x_1 < x_2 < \ldots < x_k < \ldots$; $\forall\ a > 0\ \exists\ k \in \mathbb{N}$: $x_k < a,\ x_{k+1} \geq a$. $\mathbf{M}\xi = \sum\limits_{i=1}^{+\infty} x_i p_i \geq \sum\limits_{i=k+1}^{+\infty} x_i p_i \geq a \cdot \sum\limits_{i=k+1}^{+\infty} p_i = a \cdot P\{\xi \geq a\} \Rightarrow$

$P\{\xi \geq a\} \leq \dfrac{\mathbf{M}\xi}{a}$. ■

Следствие (неравенство Чебышева): ξ – случайная величина, имеющая дисперсию; тогда $\forall\ \varepsilon > 0\ P\{|\xi - \mathbf{M}\xi| \geq \varepsilon\} \leq \dfrac{\mathbf{D}\xi}{\varepsilon^2}$.

$\triangle$ $P\{|\xi - \mathbf{M}\xi| \geq \varepsilon\} = P\{|\xi - \mathbf{M}\xi|^2 \geq \varepsilon^2\} \leq \dfrac{\mathbf{D}\xi}{\varepsilon^2}$. ■

Определение: $\eta_1, \ldots \eta_n, \ldots$ – случайные величины; последовательность $\{\eta_k\}_{k\in\mathbb{N}}$ *сходится по вероятности* к случайной величине η ($\eta_n \xrightarrow{P} \eta,\ n \to \infty$), если $\forall\ \varepsilon > 0$ $P\{|\eta_n - \eta| \geq \varepsilon\} \to 0,\ n \to \infty$ (то есть $P\{|\eta_n - \eta| < \varepsilon\} \to 1,\ n \to \infty$).

Определение: $\xi_1, \ldots \xi_n, \ldots$ – случайные величины, имеющие математическое ожидание; $S_n = \sum\limits_{k=1}^{n} \xi_k$; к последовательности $\{\xi_k\}_{k\in\mathbb{N}}$ применим *закон больших чисел*, если $\forall\ \varepsilon > 0\ P\left\{\left|\frac{S_n}{n} - \mathbf{M}\frac{S_n}{n}\right| \geq \varepsilon\right\} \to 0,\ n \to \infty$ (то есть $\left(\frac{S_n}{n} - \mathbf{M}\frac{S_n}{n}\right) \xrightarrow{P} 0,\ n \to \infty$).

Теорема 2 (Маркова): $\xi_1, \ldots \xi_n, \ldots$ – случайные величины, имеющие мат. ожидание и дисперсию. Тогда $\{\xi_k\}_{k\in\mathbb{N}}$ удовлетворяет закону больших чисел, если $\dfrac{\mathbf{D}S_n}{n^2} \to 0,\ n \to \infty$ (в случае попарной некоррелированности ξ достаточное условие примет вид $\dfrac{1}{n^2} \cdot \sum\limits_{k=1}^{n} \mathbf{D}\xi_k \to 0$, $n \to \infty$ – см. теорему 4 (3.4)).

$\triangle$ $\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}$: $\forall\ n \geq N$ $\left|\dfrac{\mathbf{D}S_n}{n^2}\right| < \varepsilon^3$, но, согласно неравенству Чебышева, $\forall\ n \geq N\ P\left\{\left|\dfrac{S_n}{n} - \mathbf{M}\dfrac{S_n}{n}\right| \geq \varepsilon\right\} \leq \dfrac{\mathbf{D}\left(\frac{S_n}{n}\right)}{\varepsilon^2} < \dfrac{\mathbf{D}S_n}{n^2\varepsilon^2} < \varepsilon \Rightarrow \left(\dfrac{S_n}{n} - \mathbf{M}\dfrac{S_n}{n}\right) \xrightarrow{P} 0,\ n \to \infty$. ■

Теорема 3 (Чебышева): $\xi_1, \ldots \xi_n, \ldots$ – попарно некоррелированные случайные величины, имеющие мат. ожидание и дисперсию; $\exists\ C \geq 0$: $\forall\ k \in \mathbb{N}\ \mathbf{D}\xi_k \leq C$. Тогда последовательность $\{\xi_k\}_{k\in\mathbb{N}}$ удовлетворяет закону больших чисел.

$\triangle$ $\dfrac{1}{n^2} \cdot \sum\limits_{k=1}^{n} \mathbf{D}\xi_k \leq \dfrac{nC}{n^2} = \dfrac{C}{n} \to 0,\ n \to \infty$, то есть последовательность $\{\xi_k\}_{k\in\mathbb{N}}$ удовлетворяет закону больших чисел по теореме Маркова. ■

Теорема 4 (Бернулли): случайные величины μ_n распределены биномиально с параметрами (n, p); тогда $\forall\ \varepsilon > 0\ P\left\{\left|\dfrac{\mu_n}{n} - p\right| < \varepsilon\right\} \to 1,\ n \to \infty$.

$\triangle$ $\mu_n = \sum\limits_{i=1}^{n} \xi_i,\ \xi_i = \begin{cases} 0,\ p \\ 1,\ q = 1 - p \end{cases}$ $\Rightarrow$ (по теореме Чебышева) $\{\xi_k\}_{k\in\mathbb{N}}$ удовлетворяет закону больших чисел $\Rightarrow \dfrac{\mu_n}{n} - \mathbf{M}\dfrac{\mu_n}{n} = \dfrac{\mu_n}{n} - p \xrightarrow{P} 0,\ n \to \infty \Rightarrow P\left\{\left|\dfrac{\mu_n}{n} - p\right| < \varepsilon\right\} \to 1$, $n \to \infty$. ■

Примеры:

1) Оценка вероятности успеха в схеме Бернулли: $P\{|\mu_n - np| \geq \varepsilon\} = \dfrac{\mathbf{D}\mu_n}{\varepsilon^2} = \dfrac{npq}{\varepsilon^2}$ ($\mathbf{M}\mu_n = np,\ \mathbf{D}\mu_n = npq$ – см. 3.4).

2) Оценка доли брака по контрольной выборке: процесс описывается гипергеометрическим распределением (белые шары – бракованные изделия); пусть n – число изделий в контрольной выборке, N – объём партии изделий, M – число бракованных изделий во всей партии, ξ – число бракованных изделий в выборке. Тогда, согласно 3.4,

$$\mathbf{M}\left(\frac{\xi}{n}\right) = \frac{M}{N},\ \mathbf{D}\left(\frac{\xi}{n}\right) = \frac{1}{n}\cdot\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1} \Rightarrow \text{(неравенство Чебышева)}$$
$$\Rightarrow P\left\{\left|\frac{\xi}{n} - \frac{M}{N}\right| \geq \delta\right\} \leq \frac{1}{n\delta^2}\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

Теорема 5 (Слуцкого – без доказательства): $g(\lambda_1, \ldots \lambda_m) \in C^1(\mathbb{R}^m)$; $\xi_n^{(1)}, \ldots \xi_n^{(m)}$ – случайные величины; $\forall\, i = \overline{1, m}\ \xi_n^{(i)} \xrightarrow{P} C_i,\ n \to \infty$. Тогда $g(\xi_n^{(1)}, \ldots \xi_n^{(m)}) \xrightarrow{P} g(C_1, \ldots C_m),\ n \to \infty$.

Определение: $\xi_1, \ldots \xi_n, \ldots$ – случайные величины; $\{\xi_k\}_{k\in\mathbb{N}}$ *сходится в среднем* порядка $r \in \mathbb{N}$ при $n \to \infty$, если $\mathbf{M}((\xi_n - n)^r) \to 0,\ n \to \infty$ (в частности, при $r = 2$ реализуется сходимость в *среднем квадратичном*).

Определение: $F_1(x), \ldots F_n(x), \ldots$ – функции распределения случайных величин $\xi_1, \ldots \xi_m, \ldots$; $\{\xi_k\}_{k\in\mathbb{N}}$ *сходится по распределению* (*слабо сходится*: $\xi_n \xrightarrow{d} \xi,\ n \to \infty$), если $\forall\, x \in \mathbb{R}$: $F \in C(x)\ \exists \lim\limits_{n\to\infty} F_n(x) = F(x)$ – функция распределения ξ.

Определение: $\xi_1, \ldots \xi_n, \ldots$ – случайные величины; если $\exists\, B > 0,\ A:\ \ \eta_n = \dfrac{\xi_n - A}{B} \xrightarrow{d} \eta$, где η имеет стандартное нормальное распределение ($F_{\eta_n}(x) \to \dfrac{1}{\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du = \Phi(x),\ n \to \infty$), то $\{\xi_k\}_{k\in\mathbb{N}}$ *асимптотически нормальна* с параметрами (A, B).

Теорема 6 (центральная предельная теорема – без доказательства): $\xi_1, \ldots, \xi_n, \ldots$ – одинаково распределённые случайные величины; $\mathbf{M}\xi_n = a,\ \mathbf{D}\xi_n = \sigma^2$; $S_n = \sum\limits_{k=1}^{n}\xi_k$. Тогда $P\left\{\dfrac{S_n - na}{\sigma\sqrt{n}} < x\right\} \underset{n\to\infty}{\overset{\mathbb{R}}{\rightrightarrows}} \Phi(x)$ (то есть S_n асимптотически нормальна с параметрами $(na, \sigma\sqrt{n})$).

Следствие: μ_n распределена биномиально с параметрами (n, p); тогда $\mu_n = \xi_1 + \ldots + \xi_n$, где ξ_i – бернуллиевские случайные величины (индикаторы). Тогда $\{\mu_n\}_{n\in\mathbb{N}}$ асимптотически нормальна с параметрами $(np, \sqrt{npq})$, то есть интегральная предельная теорема Муавра-Лапласа является следствием центральной предельной теоремы.

Теорема 7 (Ляпунова – без доказательства): $\xi_1, \ldots \xi_n, \ldots$ – случайные величины, имеющие третий центральный момент; $S_n = \sum\limits_{k=1}^{n}\xi_k,\ A_n = \sum\limits_{k=1}^{n}\mathbf{M}\xi_k,\ B_n^2 = \sum\limits_{k=1}^{n}\mathbf{D}\xi_k,\ C_n^3 = \sum\limits_{k=1}^{n}\mathbf{M}|\xi_k - \mathbf{M}\xi_k|^3$; $\dfrac{C_n}{B_n} \to 0,\ n \to \infty$; тогда $P\left\{\dfrac{S_n - \mathbf{M}S_n}{\sqrt{\mathbf{D}S_n}} < x\right\} \to \Phi(x),\ n \to \infty$.

Теорема 8 (сходимости – без доказательства): $\xi_1, \ldots \xi_n, \ldots$; $\eta_1, \ldots \eta_n, \ldots$ – случайные величины; $F_{\xi_n} \to F,\ n \to \infty$; $\eta_n \xrightarrow{P} C,\ n \to \infty$. $J_n^{(1)} = \xi_n + \eta_n,\ J_n^{(2)} = \xi_n\eta_n,\ J_n^{(3)} = \dfrac{\xi_n}{\eta_n}$. Тогда $F_{J_n^{(1)}} \to F(x - C),\ n \to \infty$; $F_{J_n^{(2)}} \to F(xC),\ n \to \infty$; $F_{J_n^{(3)}} \to F\left(\dfrac{x}{C}\right),\ n \to \infty$ (два последних соотношения верны только при $C > 0$).

4. Основы математической статистики.

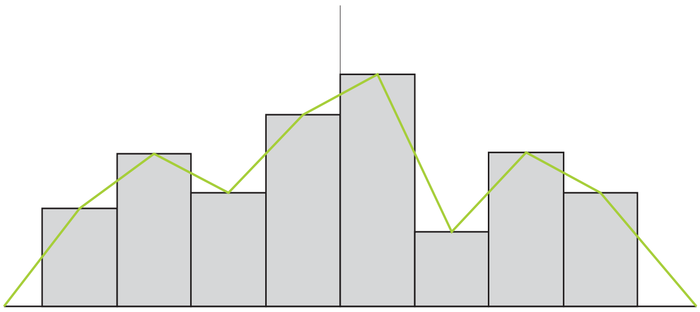
4.1. Выборки и их числовые характеристики.

Определение: $x_1, \dots x_n$ – одинаково распределённые независимые случайные величины; набор фиксированных значений $(x_1, \dots x_n)$ называется *выборкой* объёма n и обычно представляет собой результаты n одинаковых, независимых экспериментов. Элементы выборки могут быть переставлены в порядке возрастания, в результате чего получится *вариационный ряд* $(x^{(1)}, \dots x^{(n)})$, $x^{(1)} \le x^{(2)} \le \dots \le x^{(n)}$. $\widehat{F}_n(x) = \dfrac{\mu_n(x)}{n}$ называется *эмпирической функцией распределения*, где $\mu_n(x)$ – число элементов выборки, меньших x.

Теорема 1: $\forall\ x \in \mathbb{R}$, $\forall\ \varepsilon > 0$ $\lim\limits_{n\to\infty} P\{|\widehat{F}_n(x) - F(x)| \ge \varepsilon\} = 0$, где $F(x)$ – функция распределения случайных величин x_i (то есть $\widehat{F}_n \xrightarrow{P} F$, $n \to \infty$) (в данном случае $\widehat{F_n}(x)$ и $F(x)$ рассматриваются как случаные величины).

$\triangle$ При фиксированном $x \in \mathbb{R}$ μ_n имеет биномиальное распределение с параметрами (n, p) $(p = P\{x_k < x\} = F(x))$, поэтому, по теореме Бернулли, $P\left\{\left|\dfrac{\mu_n}{n} - F(x)\right| < \varepsilon\right\} \to 1$, $n \to \infty \Rightarrow \widehat{F_n} \xrightarrow{P} F$, $n \to \infty$. $\blacksquare$

Замечание: выбрав $\{z_i\}_{i=0}^{r+1} \in \overline{\mathbb{R}}$: $-\infty = z_0 < z_1 < \dots < z_r < z_{r+1} = +\infty$ и $\widehat{p}_k = \widehat{F}_n(z_{k+1}) - \widehat{F}_n(z_k)$ $(k = \overline{0, r})$, можно построить прямоугольники с основаниями на отрезках $[z_k, z_{k+1}]$ и высотой $\widehat{p}_k$; эти прямоугольники составляют *гистограмму распределения* результатов измерений случайной величины и на рисунке закрашены; линия, соединяющая середины сторон этих прямоугольников, описывает *полигон частот*, и близка к графику $p(x)$ – плотности распределения измеряемой случайной величины.



Определение: числа $a_\nu = \dfrac{1}{n} \cdot \sum\limits_{k=1}^{n} x_k^\nu$ называются *выборочными моментами*, а $m_\nu = \dfrac{1}{n} \cdot \sum\limits_{k=1}^{n} (x_k - \bar{x})^\nu$ – *центральными выборочными моментами* $(\bar{x} = a_1 = \dfrac{1}{n} \cdot \sum\limits_{k=1}^{n} x_k)$. Моменты распределённой случайной величины ξ обозначаются как $\alpha_\nu = \mathbf{M}\xi^\nu$, $\mu_\nu = \mathbf{M}(\xi - \mathbf{M}\xi)^\nu$. Очевидно, что выборочные моменты также являются случайными величинами.

Пример: рассчитаем некоторые характеристики a_1 и m_2. $\mathbf{M}\bar{x} = \dfrac{1}{n} \cdot \sum\limits_{k=1}^{n} \mathbf{M}x_k = \alpha_1$, $\mathbf{D}\bar{x} = \dfrac{1}{n^2} \cdot \sum\limits_{k=1}^{n} \mathbf{D}x_k = \dfrac{\mu_2}{n}$. Пусть $y_k = x_k - \mathbf{M}x_k$; тогда $m_2 = \dfrac{1}{n} \cdot \sum\limits_{k=1}^{n} (y_k - \bar{y})^2 = \dfrac{1}{n} \cdot \sum\limits_{k=1}^{n} y_k^2 - \bar{y}^2$

$$\left(\mathbf{M}\bar{x} = \mathbf{M}x_k \Rightarrow \mathbf{M}\bar{y} = \bar{x} - \mathbf{M}\bar{x};\ \frac{2}{n} \cdot \sum_{k=1}^{n} y_k \bar{y} = \frac{2}{n} \cdot (\bar{x} - \mathbf{M}\bar{x}) \sum_{k=1}^{n} y_k = 2\bar{y}^2\right) \Rightarrow$$

$$\Rightarrow \mathbf{M}m_2 = \frac{1}{n} \cdot \sum_{k=1}^{n} \mu_2 - \mathbf{M}(\bar{x} - \mathbf{M}\bar{x})^2 = \mu_2 - \mathbf{D}\bar{x} = \mu_2 \left(1 - \frac{1}{n}\right).$$

Определение: $\xi_0, \xi_1, \dots \xi_n$ – независимые случайные величины, имеющие стандартное нормальное распределение; тогда случайная величина $\xi_n^2 = \xi_1^2 + \ldots + \xi_n^2$ имеет *распределение хи-квадрат* с n степенями свободы, а $\tau_n = \dfrac{\xi_0 \sqrt{n}}{\chi_n^2}$ – *распределение Стьюдента* с n степенями свободы.

Теорема 2 (Фишера – без доказательства): $x_1, \dots x_n$ – независимые случайные величины, имеющие нормальное распределение с параметрами (a, σ); тогда $\bar{x}$ и m_2 независимы, причём $\bar{x}$ имеет нормальное распределение с параметрами $(a, \dfrac{\sigma}{\sqrt{n}})$, а $\dfrac{nm_2}{\sigma^2}$ – распределение хи-квадрат с $n-1$ степенями свободы.

Следствие: случайная величина $\dfrac{\bar{x}-a}{\sqrt{m_2}}\sqrt{n-1}$ имеет распределение Стьюдента с $n-1$ степенями свободы.

$\triangle$ $\dfrac{\bar{x}-a}{\sqrt{m_2}}\sqrt{n-1} = \sqrt{n-1} \cdot \dfrac{\dfrac{\bar{x}-a}{\frac{\sigma}{\sqrt{n}}}}{\dfrac{\sqrt{nm_2}}{\sigma}}$, но, по теореме Фишера, $\dfrac{\bar{x}-a}{\frac{\sigma}{\sqrt{n}}}$ имеет стандартное нормальное распределение, а $\left(\dfrac{\sqrt{nm_2}}{\sigma}\right)^2$ – распределение хи-квадрат с $n-1$ степенями свободы. ■

4.2. Точечные и интервальные оценки.

Определение: *статистикой* называется любая функция выборки. *Точечная оценка* параметра θ – функция $\widehat{\theta}(x_1, \dots x_n)$, где $(x_1, \dots x_n)$ – выборка. Точечная оценка называется *несмещённой*, если $\mathbf{M}\widehat{\theta} = \theta$ и *состоятельной*, если $\widehat{\theta}_n \xrightarrow{P} \theta, \ n \to \infty$.

Определение: *интервальной оценкой* параметра θ называются функции $\underline{\theta}(x_1, \dots x_n)$ и $\overline{\theta}(x_1, \dots x_n)$: $\forall \ (x_1, \dots x_n) \ \underline{\theta}(x_1, \dots x_n) < \theta < \overline{\theta}(x_1, \dots x_n)$. Если $P\{\underline{\theta}(x_1, \dots x_n) < \theta < \overline{\theta}(x_1, \dots x_n)\} = 1 - 2\alpha$, то $[\underline{\theta}, \overline{\theta}]$ называется *доверительными интервалом* для θ, соответствующим доверительной вероятности $1 - 2\alpha$.

Способы получения оценок:

1. *Непосредственный подбор* (критерием правильности является близость a_ν и α_ν).

2. *По наибольшему правдоподобию*: введём *функцию правдоподобия* $L(x_1, \dots x_n, \theta) = p_{x_1}(x_1, \theta) \cdot \ldots \cdot p_{x_n}(x_n, \theta)$ ($x_1, \dots x_n$ – разные случайные величины, а $p_1, \dots p_n$ – их плотности распределения); оценка для θ выбирается так, чтобы значение функции L было максимальным, то есть из уравнения $\dfrac{\partial L}{\partial \theta} = 0$ (или $\dfrac{\partial \ln L}{\partial \theta} = 0$).

3. *С помощью критерия* χ^2: при оценке s параметров $(\theta_1, \dots \theta_s)$ распределение величины $\sum\limits_{k=1}^{n} \dfrac{(x_k - np_k(\theta_1, \dots \theta_s))^2}{np_k(\theta_1, \dots \theta_s)(1 - p_k(\theta_1, \dots \theta_s))}$ близко к χ^2_{n+s-1}; зная это, можно определить функции $p_k(\theta_1, \dots \theta_s)$ $(k = \overline{1, n})$, а с их помощью – точечные оценки для $\theta_1, \dots \theta_s$.

Доверительные интервалы для нормального распределения:

1. *Оценка a при известном σ*: по теореме Фишера $\dfrac{\bar{x}-a}{\frac{\sigma}{\sqrt{n}}}$ имеет стандартное нормальное распределение, то есть $\exists \ u_\alpha > 0$: $P\left\{\left|\dfrac{\bar{x}-a}{\frac{\sigma}{\sqrt{n}}}\right| < u_\alpha\right\} = 1 - 2\alpha$, u_α определяется из уравнения $\dfrac{1}{\sqrt{2\pi}} \cdot \int\limits_{u_\alpha}^{+\infty} e^{-\frac{x^2}{2}} dx = \alpha \Rightarrow P\left\{\bar{x} - u_\alpha \cdot \dfrac{\sigma}{\sqrt{n}} < a < \bar{x} + u_\alpha \cdot \dfrac{\sigma}{\sqrt{n}}\right\} = 1 - 2\alpha$, то есть

$\left[\bar{x} - u_\alpha \cdot \dfrac{\sigma}{\sqrt{n}},\ \bar{x} + u_\alpha \cdot \dfrac{\sigma}{\sqrt{n}}\right]$ – доверительный интервал для a.

2. *Оценка для a при неизвестном σ*: согласно следствию из теоремы Фишера $\tau_{n-1} = \dfrac{\bar{x} - a}{\sqrt{m_2}}\sqrt{n-1}$ имеет распределение Стьюдента с $n-1$ степенями свободы; значит, $\exists\ t_{\alpha,n-1}$: $P\{|\tau_{n-1}| < t_{\alpha,n-1}\} = 1 - 2\alpha$, поэтому $P\left\{\bar{x} - t_{\alpha,n-1} \cdot \sqrt{\dfrac{m_2}{n-1}} < a < \bar{x} + t_{\alpha,n-1} \cdot \sqrt{\dfrac{m_2}{n-1}}\right\} = 1 - 2\alpha$.

3. *Оценка для σ при известном a*: $S^2 = \sum\limits_{k=1}^{n}(x_k - a)^2 \Rightarrow \dfrac{S^2}{\sigma^2}$ имеет распределение хи-квадрат с n степенями свободы. $\forall\ \alpha \in [0, \dfrac{1}{2}\ ]\ \exists\ \chi_{\alpha,n}$: $P\{\chi^2_n > \chi^2_{\alpha,n}\} = \alpha,\ \exists \chi_{1-\alpha,n}$: $P\{\chi^2_n > \chi^2_{1-\alpha,n}\} = 1 - \alpha \Rightarrow$

$$P\left\{\chi^2_{1-\alpha,n} < \frac{S^2}{\sigma^2} < \chi^2_{\alpha,n}\right\} = 1 - 2\alpha \Rightarrow P\left\{\frac{S}{\chi_{\alpha,n}} < \sigma < \frac{S}{\chi_{1-\alpha,n}}\right\} = 1 - 2\alpha.$$

4. *Оценка для σ при неизвестном a*: по теореме Фишера $\dfrac{nm_2}{\sigma^2}$ имеет распределение хи-квадрат с $n-1$ степенями свободы, то есть $P\left\{\chi^2_{1-\alpha,n-1} < \dfrac{nm_2}{\sigma^2} < \chi^2_{\alpha,n-1}\right\} = 1 - 2\alpha \Rightarrow$ $P\left\{\dfrac{\sqrt{nm_2}}{\chi_{\alpha,n-1}} < \sigma < \dfrac{\sqrt{nm_2}}{\chi_{1-\alpha,n-1}}\right\} = 1 - 2\alpha$.

5. *Сравнение выборок*: $(x_{11}, \dots x_{n_1 1})$, $(x_{12}, \dots x_{n_2 2})$ – независимые выборки; $\bar{x}_i = \dfrac{1}{n_i} \cdot \sum\limits_{k=1}^{n_i} x_{ki}$, $m_{2i} = \dfrac{1}{n_i} \cdot \sum\limits_{k=1}^{n_i}(x_{ki} - \bar{x}_i)^2$ $(i = \overline{1,2})$; тогда, по теореме Фишера, $\chi^2_{n_1+n_2-2} = \dfrac{n_1 m_{21} + n_2 m_{22}}{\sigma^2}$ имеет распределение хи-квадрат с $n_1 + n_2 - 2$ степенями свободы.

$$\tau_{n_1+n_2-2} = \frac{\dfrac{\bar{x}_1 - \bar{x}_2 - (a_1 - a_2)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}}{\sqrt{\dfrac{n_1 m_{21} + n_2 m_{22}}{(n_1 + n_2 - 2)\sigma^2}}} = \frac{(\bar{x}_1 - \bar{x}_2 - (a_1 - a_2))\sqrt{n_1 n_2 (n_1 + n_2 - 2)}}{\sqrt{(n_1 m_{21} + n_2 m_{22})(n_1 + n_2)}} \Rightarrow$$

$$P\left\{|\bar{x}_1 - \bar{x}_2| - t_{\alpha,n_1+n_2-2} \cdot \sqrt{\frac{(n_1 m_{21} + n_2 m_{22})(n_1 + n_2)}{n_1 n_2 (n_1 + n_2 - 2)}} < |a_1 - a_2| < \right.$$
$$\left. < |\bar{x}_1 - \bar{x}_2| + t_{\alpha,n_1+n_2-2} \cdot \sqrt{\frac{(n_1 m_{21} + n_2 m_{22})(n_1 + n_2)}{n_1 n_2 (n_1 + n_2 - 2)}}\right\} = 1 - 2\alpha.$$

Определив доверительный интервал, можно сделать предположение о принадлежности двух выборок к одной и той же или разным генеральным совокупностям.

Замечание: если случайные величины $x_1, \dots x_n$ распределены так, что имеют дисперсию ($\mathbf{M}x_k = a$, $\mathbf{D}x_k = \sigma^2$), то, согласно центральной предельной теореме, распределение $\dfrac{\bar{x} - a}{\frac{\sigma}{\sqrt{n}}}$ близко к стандартному нормальному при больших n. Это позволяет оценивать a и σ с помощью полученных формул. Например, для биномиального распределения $P\left\{\dfrac{\mu_n}{n} - u_\alpha \cdot \dfrac{\sigma_n}{\sqrt{n}} < p < \dfrac{\mu_n}{n} + u_\alpha \cdot \dfrac{\sigma_n}{\sqrt{n}}\right\} \to 1 - 2\alpha,\ n \to \infty$, где σ^2_n – любая состоятельная оценка σ^2.

4.3. Статистическая проверка гипотез.

1. *С помощью интервальных оценок*: случайная величина ξ распределена нормально с параметрами (a, σ); гипотеза состоит в том, что $\mathbf{M}\xi = a_0$; из 4.2

$$P\left\{|\bar{x} - a_0| > t_{\alpha,n-1} \cdot \sqrt{\frac{m_2}{n-1}}\right\} = 2\alpha.$$

Если, согласно измерениям, $|\bar{x} - a_0| > t_{\alpha,n-1} \cdot \sqrt{\dfrac{m_2}{n-1}}$, то гипотеза отвергается; в этом случае вероятность отказа от верной гипотезы равна 2α.

2. *Критерий хи-квадрат*: гипотеза утверждает, что случайная величина x_k имеет функцию распределения $F(x)$. Выберем произвольные числа $z_1, \dots z_r$: $z_1 < \dots < z_r$. Если гипотеза верна, то $p_l = P\{x_k \in [z_l, z_{l+1})\} = F(z_{l+1}) - F(z_l)$. Пусть m_l – число результатов эксперимента, попавших в отрезок $[z_l, z_{l+1})$; в случае верного предположения $\mathbf{M}m_l = np_l$; мерой расхождения является $\eta_{n,r} = \sum\limits_{i=1}^{r+1} \dfrac{(m_i - np_i)^2}{np_i(1-p_i)}$, распределение которой близко к распределению хи-квадрат с r степенями свободы. Поэтому $P\{\eta_{n,r} < C\} \to P\{\chi_r^2 < C\}$, $n \to \infty$, где C определяется доверительной вероятностью, которая равна вероятности отказа от верной гипотезы.

3. *Критерий Колмогорова*:

Теорема 1 (Колмогорова – без доказательства):

$$P\left\{\sqrt{n} \cdot \sup_{x \in \mathbb{R}} |\widehat{F}_n(x) - F(x)| < z\right\} \to K(z),\ n \to \infty,\ \text{где}$$

$$K(z) = \begin{cases} 0,\ z \le 0 \\ \sum\limits_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2 z^2},\ z > 0. \end{cases}$$

Теорема позволяет определить z_α: $1 - K(z_\alpha) = \alpha$; тогда

$$P\left\{\widehat{F}_n(x) - \frac{z_\alpha}{\sqrt{n}} < F(x) < \widehat{F}_n(x) + \frac{z_\alpha}{\sqrt{n}}\right\} \to 1 - \alpha,\ n \to \infty.$$

Определение: *статистикой критерия* называется функция, значение которой определяет принятие гипотезы или отказ от неё (например, $\eta_{n,r}$ для критерия хи-квадрат). *Критической областью* называется область $\mathbb{R}^n$, в которой статистика критерия превышает величину, заданную доверительной вероятностью.

Определение: гипотеза H *проста*, если она однозначно определяет распределение выборки, и *сложна* в противном случае.

Определение: пусть H_0, H_1 – простые конкурирующие гипотезы; S – критическая область. *Ошибкой первого рода* $\alpha = P_0\{x \in S\}$ называется вероятность отвергнуть гипотезу H_0 в том, случае, когда она верна; *ошибкой второго рода* $\beta = P_1\{x \notin S\}$ называется вероятность принять гипотезу H_0 в том случае, когда она ложна. Число $1 - \beta$ называют *мощностью критерия.*

Теорема 2 (критерий Неймана-Пирсона – без доказательства): пусть гипотезы H_0 и H_1 задают функции $p_0(\overrightarrow{x})$ и $p_1(\overrightarrow{x})$; $S_c = \{\overrightarrow{x} | p_1(\overrightarrow{x}) \ge cp_0(\overrightarrow{x})\}$. Пусть $\forall\ \alpha \in [0; 1]\ \exists\ c$: $P_0\{\overrightarrow{x} \in S_c\} = \alpha$; тогда наиболее мощным (при фиксированном α) является критерий, определяемый областью S_c.

4.4. Метод наименьших квадратов.

Пусть задан набор результатов измерений (точек) $(x_1, y_1), \ldots (x_n, y_n)$, который требуется аппроксимировать (приблизить) линейной функцией $y = ax + b$, причём $\forall\ i = \overline{1, n}\ y_i = ax_i + b + \delta_i$, где δ_i независимы и распределены нормально с параметрами $(0, \sigma)$. Воспользуемся критерием правдоподобия: $L(y, a, b, \sigma^2) = \dfrac{1}{(\sqrt{2\pi}\sigma)^n} \cdot \exp\left(-\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(y_i - ax_i - b_i)^2\right)$. Тогда

$$\frac{\partial \ln L}{\partial a} = \frac{2}{\sigma^2} \cdot \sum_{i=1}^{n} x_i(y_i - ax_i - b) = 0,$$

$$\frac{\partial \ln L}{\partial b} = \frac{2}{\sigma^2} \cdot \sum_{i=1}^{n} (y_i - ax_i - b) = 0,$$

$$\frac{\partial \ln L}{\partial (\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{\sigma^4} \cdot \sum_{i=1}^{n} (y_i - ax_i - b)^2 = 0.$$

Выберем x_i: $\sum\limits_{i=1}^{n} x_i = 0$; в этом случае

$$a^* = \frac{\sum\limits_{k=1}^{n} x_i y_i}{\sum\limits_{i=1}^{n} x_i^2},\ b^* = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i,\ {\sigma^*}^2 = \frac{2}{n} \cdot \sum_{i=1}^{n} (y_i - a^* x_i - b^*)^2.$$

Подставляя $y_i = ax_i + b + \delta_i$, получим

$$a^* = a + \frac{b\sum\limits_{i=1}^{n} x_i}{\sum\limits_{i=1}^{n} x_i^2} + \frac{\sum\limits_{i=1}^{n} x_i \delta_i}{\sum\limits_{i=1}^{n} x_i^2},\ b^* = b + a\sum_{i=1}^{n} x_i + \frac{1}{n} \cdot \sum_{i=1}^{n} \delta_i \Rightarrow \mathbf{M}a^* = a,\ \mathbf{M}b^* = b,$$

то есть полученные оценки являются несмещёнными.

Тот же результат может быть получен при минимизации функции $Q(a, b) = \sum\limits_{i=1}^{n} (y_i - ax_i - b)^2$, поэтому такой метод аппроксимации называют *методом наименьших квадратов* (МНК).

Предметный указатель