Теория вероятности. Теория

Александр Сергеев

1 Вероятностное пространство. Вероятность и ее свойство

Определение
Алгебра событий:
Ω – множество элементарных исходов
$\mathcal{A}$ – набор подмножеств Ω
$\mathcal{A}$ – алгебра, если

1. $\Omega \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow \overline{A} = \Omega \setminus A \in \mathcal{A}$
3. $A, B \in \mathcal{A} \Rightarrow A \cup B = A + B \in \mathcal{A}$

Элементы алегбры – *события*
Операции с событиями

1. $A \cup B = A + B$
2. $A \cap B = AB = \overline{A} + \overline{B}$
3. $\overline{A} = \Omega \setminus A$
4. $A \setminus B = A - B = A\overline{B}$

Определение
σ-алгебра
$\mathcal{A}$ – сигма-алгебра

1. $\mathcal{A}$ – алгебра

2. $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup A_i \in \mathcal{A}$

Определение
События A, B – *несовместные* $AB = \varnothing$
Набор несовместный, если события попарно несовместные
Определение(вероятностное пространство)
Ω – множество элементарных исходов
$\mathcal{A}$ – сигма-алгебра
$P : \mathcal{A} \to \mathbb{R}$ – вероятность, если

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. $P(\bigcup A_i) = \sum P(A_i)$

Определение(вероятностное пространство в широком смысле)
Ω – множество элементарных исходов
$\mathcal{A}$ – алгебра
$P : \mathcal{A} \to \mathbb{R}$ – вероятность, если

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. $\bigcup^{\infty} A_i \Rightarrow P(\bigcup^{\infty} A_i) = \sum P(A_i)$

Теорема о продолжении меры
$\langle \Omega, \mathcal{A}, P \rangle$ – вероятностное пространство в широком смысле
Тогда $\exists! Q : \sigma(\mathcal{A}) \to \mathbb{R}$ – вероятность, $Q\Big|_{\mathcal{A}} = P$, где $\sigma(\mathcal{A})$ – сигма-алгебра, содержащая $\mathcal{A}$
Определение
$\mathcal{A}$ – система интервалов на $\mathbb{R}$, замкнутая относительно конечного объединения и пересечения
$\mathcal{B} = \sigma(\mathcal{A})$ – борелевская сигма-алгебра
Примеры вероятностных пространств

1. Модель классической вероятности
   $\Omega = \{\omega_1, \dots, \omega_N\}$
   $\mathcal{A} = 2^\Omega$
   $P(\{\omega_i\}) = P(\omega_i) = \dfrac{1}{N}$
   $\mathcal{A} = \{\omega_{i_1}, \dots, \omega_{i_M}\} \Rightarrow P(\mathcal{A}) = \dfrac{M}{N}$

2. Ω – набор $\{0^i, 1\}, i \in \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$
   $P(0^i 1) = q^i p$

3. Модель геометрической вероятности Ω – ограниченное, измеримое по Лебегу множество
   $\mathcal{A}$ – измеримое по Лебегу подмножество Ω
   $P(A) = \dfrac{\lambda A}{\lambda \Omega}$

Теорема (свойство вероятности)

1. $A \subset B \Rightarrow P(A) \leq P(B)$
2. $P(A) \leq 1$
3. $P(A) + P(\overline{A}) = 1$
4. $P(A + B) = P(A) + P(B) - P(AB)$
5. $P(\varnothing) = 0$
6. $P(\bigcup A_i) \leq \sum P(A_i)$

Доказательство

1. $P(B) = P(A) + P(B - A)$
2. $A \subset \Omega \Rightarrow P(A) \leq 1$
3. $A \sqcup \overline{A} = \Omega$
4. $B = AB \sqcup (B \setminus AB)$
5. $B_1 = A_1$ $B_2 = A_2 \setminus A_1$
   $B_n = A_n \setminus (A_1 \cup \dots A_{n-1}) \bigsqcup B_i = \bigcup A_i$
   $B_i \subset A_i$
   $P(\bigcup A_i) = P(\bigsqcup B_i) = \sum P(B_i) \leq \sum P(A_i)$

Теорема (формула включения/исключения)
$$P(A_1+\ldots+A_n)=\sum_i P(A_i)-\sum_{i<j}P(A_iA_j)+\sum_{i<j<k}P(A_iA_jA_k)+\ldots(-1)^{n+1}\sum_{i_1<\ldots<i_n}P(A_{i_1}\ldots A_{i_n})$$
Доказательство
Доказательство по индукции
Теорема
$\mathcal{A}$ – алгебра(?) на Ω, p – мера
Тогда равносильны

1. p – счетно-аддитивно
2. p – конечно-аддитивно$+\forall\,(B_n)_{n=1}^{\infty}: B_{n+1}\subset B_n, B=\bigcap B_n \Rightarrow P(B_n)\to P(B)$ – непрерывность сверху
3. p – конечно-аддитивно$+\forall\,(A_n)_{n=1}^{\infty}: A_{n+1}\supset A_n, A=\bigcup A_n \Rightarrow P(A_n)\to P(A)$ – непрерывность снизу

Доказательство (непрерывность сверху) $\Leftrightarrow$ (непрерывность снизу)
$A(n): A_n\subset A_{n+1}; A=\bigcup A_n$
$B_n:=\overline{A}_n, B:=\overline{A}$
$B_n\supset B_{n+1}$
$B=\overline{A}=\overline{\bigcup A_n}=\cap\overline{A}_n=\cap B_n$
$p(B_n)=1-p(A_n)\to 1-p(A)=p(B)$
Доказательство $1\to 2$
$C_1=B_1\overline{B}_2$
$C_2=B_2\overline{B}_3$
$C_k=B_k\overline{B}_{k+1}$
$$B_k=B\sqcup\bigsqcup_{j=k}^{\infty}C_j$$
$$p(B_k)=p(B)+\underbrace{\sum_{j=k}^{\infty}p(C_j)}_{\to 0}$$
$p(B_k)\to p(B)$
Доказательство $2\to 1$
$$\sum_{k=1}^{\infty}p(C_k)=\lim_n\sum_{k=1}^{n}p(C_k)=\lim_n p(\bigsqcup_{k=1}^{n}C_k)=p(B)$$

2 Условная вероятность. Формуоа полной вероятности. Теорема Байеса

Определение (условная вероятность)
$(\Omega, \mathcal{A}, p)$ – вероятностное пространство
$B \in \mathcal{A} : p(B) > 0$
$p_B(A) = p(A|B) := \dfrac{p(AB)}{p(B)}$
Замечание
$p(AB) = p(A)p(B|A)$
Теорема (формула произведения вероятностей)
$p(A_1 \dots A_1) = p(A_1)p(A_2|A_1) \dots p(A_n|A_1 \dots A_{n-1})$
Доказательство
Тривиально
Определение
$A_1, \dots, A_n$ – независимые, если $\forall\, A_{i_1}, \dots, A_{i_k}\ p(A_{i_1} \dots A_{i_k}) = p(A_{i_1}) \dots p(A_{i_k})$
Теорема (формула полной вероятности)
$A \subset \bigsqcup\limits_k B_k$ (как правило, $\bigsqcup B_k = \Omega$)
Тогда $p(A) = \sum\limits_k p(A|B_k)p(B_k)$
Доказательство
$p(A) = p(A \cap \bigsqcup B_k) = p(\bigsqcup AB_k) = \sum\limits_k p(AB_k) = \sum p(A|B_k)p(B_k)$
Теорема Байеса
Краткая форма:
$p(B|A) = \dfrac{p(A|B)p(B)}{p(A)}$
Полная:
$A \subset \bigsqcup B_k$

$$\underbrace{p(B_k|A)}_{\text{апостериорные; posterior}} = \frac{\underbrace{p(A|B_k)}_{\text{likelyhood}}\ \underbrace{p(B_k)}_{\text{априорные; prior}}}{\sum_j p(A|B_j)} p(B_j)$$

Доказательство краткой формы
$p(B|A) = \dfrac{p(AB)}{p(A)} = \dfrac{p(A|B)p(B)}{p(A)}$

3 Схема Бернулли. Полиноминальная схема. Предельные теоремы, связь со схемой Бернулли

Определение (независимые испытания)
$n \in \mathbb{N}$
$(\Omega_1, \mathcal{A}_1, p_1), \ldots (\Omega_n, \mathcal{A}_n, p_n)$ – вероятностные пространтсва, описывающие виды экспериментов
$\Omega^{(n)} = \Omega_1 \times \ldots \times \Omega_n$
$\mathcal{A}^{(n)} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_n$
$p^{(n)} : \mathcal{A}^{(n)} \to \mathbb{R}$
$p^{(n)}(A_1 \times \ldots \times A_n) = p_1(A_1) \cdot \ldots \cdot p_n(A_n)$
$(\Omega^{(n)}, \mathcal{A}^{(n)}, p^{(n)})$ – описание n независимых испытаний
Замечание
$\mathcal{A}^{(n)}$ может не быть сигма-алгеброй
$(\Omega^{(n)}, \mathcal{A}^{(n)}, p^{(n)})$ – вероятностное пространство в широком смысле
$\sigma(\mathcal{A}^{(n)})$ – сигма-алгебра
Определение (схемы Бернулли)
$n \in \mathbb{N}$
$p \in (0, 1)$ – вероятность успеха
$q = 1 - p$ – вероятность неудачи
$\Omega = \{0, 1\}$
$\mathcal{A} = 2^\Omega$
$\Omega^{(n)} = \{0, 1\}^n$
$w \in \Omega^{(n)}$ – вектор из нуля и единиц
$p(w) = p^{\sum w_i} q^{n - \sum w_i}$
$(\Omega^{(n)}, \mathcal{A}^{(n)}, p^{(n)})$ – схема Бернулли

Теорема
S_n – количество успехов в n испытаниях
$p(S_n = k) = \binom{n}{k} p^k q^{n-k}$
Теорема (о наиболее вероятном k_*)
$p_k := p(S_n = k)$
$\frac{p_k}{p_{k+1}} = \frac{(k+1)q}{(n-k)p} \vee 1$
$(k+1)(1-p) \vee (n-k)p$

$k+1-pk-p \vee np-pk$
$k \vee (n+1)p-1$

1. $p(n+1) \in \mathbb{N}$
   Тогда $\exists\, k_0 : k_0 = p(n+1)-1$
   Если $k < k_0$, то $p_k < p_{k_0+1}$
   Если $k = k_0$, то $p_{k_0} = p_{k_0+1}$
   Если $k > k_0$, то $p_k > p_{k_0+1}$
   Т.о. k_0, k_0+1 – наиболее вероятные

2. $p(n+1) \notin \mathbb{N}$
   Тогда $\exists\, k_1 : p_{k_1} < p_{k_1+1}, p_{k_1+1} > p_{k_1+2}$
   Т.о. k_1 – наиболее вероятные

Тогда $k_* = \begin{cases} \lceil p(n+1)-1 \rceil, & p(n+1) \notin \mathbb{N} \\ p(n+1)-1, p(n+1), & p(n+1) \in \mathbb{N} \end{cases}$
Утверждение
$p(S_n \geq 1) = 1-q^n$
Утверждение
$\alpha \in (0,1)$
$p \in (0,1)$
$n : p(S_n \geq 1) \geq \alpha$
Ответ: $n = \lceil \log_q(1-\alpha) \rceil$
Определение (полиноминальная схема)
n – количество испытаний
m – количество возможных исходов
$p = (p_1, \ldots, p_m)$ – вектор вероятностей исходов
$\sum p = 1$
$\Omega = \{(i_1, \ldots, i_m)^T \in \{0,1\}^m : \sum\limits_j i_j = 1\}$ – множество столбцов с одной единицей
$\Omega^{(n)} = \{A \in M_{m\times n} : A_{ij} \in \{0,1\}, \sum\limits_i A_{ij} = 1\}$
$p(A) = p_1^{\sum_j A_{1,j}} \cdot \ldots \cdot p_m^{\sum_j A_{m,j}}$
Теорема
$S_{n,j}$ – количество исходов типа j в n испытаниях
$p(S_{n,1} = k_1, \ldots, S_{n,m} = k_m) = \dfrac{n!}{k_1! \ldots k_m!} p_1^{k_1} \ldots p_m^{k_m}, \sum k_j = n$
Теорема (закон больших чисел в форме Бернулли)

$\forall \varepsilon \ p(|\frac{S_n}{n} - p| > \varepsilon) \underset{n}{\to} 0$

Теорема (теорема Пуассона)

$p_n = \frac{\lambda}{n} + o(\frac{1}{n}), n \to \infty$ – схема серий из n испытаний (вероятность успеха при n испытаниях)

Тогда $p(S_n = k) \xrightarrow[n\to\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}$

Доказательство

$p(S_n = k) = \frac{n!}{k!(n-k)!}(\frac{\lambda}{n} + o(\frac{1}{n}))^k(1 - \frac{\lambda}{n} + o(\frac{1}{n}))^{n-k} =$

$= \frac{1}{k!}\frac{\prod_{j=0}^{k-1}(n-j)}{n^k}(\lambda + o(\frac{1}{n}))^k\frac{(1-\frac{\lambda}{n}+o(\frac{1}{n}))^n}{(1-\frac{\lambda}{n}+o(\frac{1}{n}))^k} = \frac{\lambda^k}{k!}e^{-\lambda}$

Замечание (о погрешности)

$\lambda = np$

Тогда $|p(S_n = k) - e^{-\lambda}\frac{\lambda^k}{k!}| \leq np^2 = \frac{\lambda^2}{n}$

Лемма 1

$k \to \infty, (n \geq k \Rightarrow n \to \infty)$

$p_* = \frac{k}{n}$

$H(x) = x \ln\frac{x}{p} + (1-x)\ln\frac{1-x}{1-p}$ – энтропия

$n - k \to \infty \Rightarrow p(S_n = k) \sim \frac{1}{\sqrt{2\pi n p_*(1-p_*)}}\exp(-nH(p_*))$

Доказательство

$p(S_n = k) = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \sim \frac{\sqrt{2\pi n}n^n e^{-n}p^k(1-p)^{n-k}}{\sqrt{2\pi k}k^k e^{-k}\sqrt{2\pi(n-k)}(n-k)^{n-k}e^{-(n-k)}} =$

$\frac{n^k n^{n-k}(1-p)^{n-k}p^k}{\sqrt{2\pi p_* n(1-p_*)}k^k(n-k)^{n-k}} = \frac{1}{\sqrt{\cdots}}\exp\ln\frac{n^k p^k}{k^k}\cdot\frac{n^{n-k}(1-p)^{n-k}}{(n-k)^{n-k}}$

$= \frac{1}{\sqrt{\cdots}}\exp k\ln\frac{p}{p_*} + (n-k)\ln\frac{n(1-p)}{n-k}$

$= \frac{1}{\sqrt{\cdots}}\exp(-n\underbrace{(p_*\ln\frac{p}{p_*} + (1-p_*)\ln\frac{1-p_*}{1-p})}_{H(p_*)})$

Лемма 2

$H(p_*) = \frac{1}{2p(1-p)}(p-p_*)^2 + O((p-p_*)^3)$

При $p_* \to p$

Доказательство

$H(p) = 0$

$H'(x) = \ln \dfrac{x}{p} - 1 - \ln \dfrac{1-x}{1-p} + 1; H'(p) = 0$

$H''(x) = \dfrac{1}{x} + \dfrac{1}{1-x} = \dfrac{1}{x(1-x)}; H''(p) = \dfrac{1}{p(1-p)}$

Теорема (локальная предельная Муавра-Лапласа)

Требуем Лемму 1 + $(k - np = o(n^{\frac{2}{3}}))$

Тогда $p(S_n = k) \sim \dfrac{1}{\sqrt{2\pi np(1-p)}} \exp(-\dfrac{(k-np)^2}{2np(1-p)})$

Доказательство

$p_* - p = o(n^{-\frac{1}{3}})$ – из $k - np = o(n^{\frac{2}{3}})$

Применим Л2

$$\underbrace{\frac{1}{\sqrt{2\pi np_*(1-p_*)}} \exp(-nH(p_*))}_{\text{из Л1}} = \frac{1}{\sqrt{2\pi np_*(1-p_*)}} \exp -n(\frac{(p-p_*)^2}{2p(1-p)} + O((p-p_*)^3)) \sim$$

$$\frac{1}{\sqrt{2\pi np(1-p)}} \exp(-\frac{n(\frac{k}{n}-p)^2}{2p(1-p)} + nO(o(n^{-1}))) =$$

$$\frac{1}{\sqrt{2\pi np(1-p)}} \exp(-\frac{(k-np)^2}{2np(1-p)} + o(1))$$

Теорема(интегральная Муавра-Лапласа)

$F_n(x) = p(\dfrac{S_n - np}{\sqrt{npq}} \leq x)$

$\Phi(x) = \displaystyle\int_{-\infty}^{x} \phi$

$\phi = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Тогда $\displaystyle\sup_{-\infty \leq x_1 \leq x_2 \leq +\infty} |F_n(x_2) - F_n(x_1) - \underbrace{(\Phi(x_2) - \Phi(x_1))}_{\frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} \,\mathrm{d}\, t}| \xrightarrow[n\to\infty]{} 0$

Замечание

$\displaystyle\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \leq C\frac{p(1-p)^3 + (1-p)p^3}{(pq)^{\frac{3}{2}}\sqrt{n}} = C\frac{(1-p)p((1-p)^2+p^2)}{(pq)^{\frac{3}{2}}\sqrt{n}} \leq$

$C\dfrac{1}{\sqrt{pqn}}$

По последним оценкам $C < 1$

4 Случайные величины. Распределение случайной величины

Определение
Пусть $(\Omega, \mathcal{A}, P)$ – вероятностное пространство
$\mathcal{B}$ – борелевская сигма-алгебра (минимальная сигма-алгебра, содержащая все интервалы)
$X : \Omega \to \mathbb{R}$ – случайная величина, если X – измеримо, т.е. $\forall B \in \mathcal{B}$ – борелевская сигма-алгебра $X^{-1}(B) \in \mathcal{A}$
$P_X : \mathcal{B} \to \mathbb{R}$ – распределение случайной величины, если $P_x(B) = P(\{\omega : X(\omega) \in B\})$ для всех $B \in \mathbb{B}$
Замечание (обозначеиня)
Случайные величины обозначаются большими латинскими буквами из конца алфавита(X, Y, U, W) или маленькими греческими ξ, ν, η
Замечание
P_X удовлетворяет аксиомам вероятности
Определение
Функция распределенич $F_X(t) = P(X \leq t) = P(\{\omega : X(\Omega) \leq t\}), t \in \mathbb{R}$
Теорема (свойства функции распределения)

1. F не убывает
2. $F(+\infty) = 1, F(-\infty) = 0$
3. F – непрерывна справа

Доказательство

1. $t_1 < t_2$
   $F(t_2) = F(t_1) + P(\{\omega : t_1 < X(\Omega) \leq t_2\}) \geq F(t_1)$
2. Пусть t_n – монотонно возрастает к $+\infty$
   $(-\infty, t_n] \subset (-\infty, t_{n+1}]$
   $\bigcup(-\infty, t_n] = \mathbb{R}$
   Тогда $F(t_n) \to P(X \in \mathbb{R}) = 1$
   Пусть t_n – монотонно убывает к $-\infty$

$(-\infty, t_n] \supset (-\infty, t_{n+1}]$
$\bigcap (-\infty, t_n] = \varnothing$
Тогда $F(t_n) \to P(X \in \varnothing) = 0$

3. Пусть t_n – монотонно стремится к 0
$F(x_0 + t_n) \to F(x_0)$
$(-\infty, x_0 + t_{n+1}] \subset (-\infty, x_0 + t_n]$
$\bigcap (-\infty, x_0 + t_n] = (-\infty, x_0]$
$F(x_0 + t_n) = P(X \leq x_0 + t_n) \to P(X \leq x_0) = F(x_0)$

Замечание (непрерывность слева)
Пусть t_n – монотонно стремится к 0
$F(x_0) - F(x_0 - t_n) = P(X \leq x_0) - P(X \leq x_0 - t_n) = P(x_0 - t_n < X \leq x_0)$
$(x_0 - t_n, x_0] \supset (x_0 - t_{n+1}, x_0]$
$\bigcap (x_0 - t_n, x_0] = \{x_0\}$
Т.о. $F(x_0) - F(x_0 - t_n) \to P(X = x_0)$ – иногда не 0
Т.о. нет непрерывности слева
Лемма
$\forall (B_n) : B_{n+1} \subset B_n, B = \bigcap B_n \ P(B_n) \to P(B)$
Равносильно $\forall (A_n) : A_{n+1} \subset A_n, \bigcap A_n = \varnothing \ P(A_n) \to 0$
Доказательство $\Leftarrow$
$A_n = B_n \setminus B$
$A_{n+1} \subset A_n$
$\bigcap A_n = \varnothing$
$\underbrace{P(A_n)}_{P(B_n) - P(B)} \to 0$
Теорема (о достаточности F для описания вероятностного распределения)
Пусть F не убывает, $F(+\infty) = 1, F(-\infty) = 0, F$ – непрерывно справа
Тогда $\exists (\Omega, \mathcal{A}, P), X$ – случайная величина, такие что $F_X = F$
Доказательство
$\Omega := \mathbb{R}$
$\mathcal{A}$ – система интервалов вида $(a, b]$ (+все лучи и $\mathbb{R}$), замкнутая относительно конечного числа $\sqcup$, т.е. $\mathcal{A}$ – алгебра
$A := \bigsqcup\limits_{k=1}^{n} (a_k, b_k], P(A) := \sum\limits_{k=1}^{n} (F(b_k) - F(a_k))$

$P(\mathbb{R}) = F(+\infty) - F(-\infty) = 1$
Проверим счетную аддитивность
Пусть $(A_n) : A_{n+1} \subset A_n, \bigcap A_n = \varnothing$

$A_n = \bigsqcup\limits_{k=1}^{k_n} (a_{nk}, b_{nk}]$
Пусть все $A_n \subset [-M, M]$
Тогда $\exists\, (B_n) : cl(B_n) \subset A_n$ и $P(A_n) - P(B_n) < \dfrac{\varepsilon}{2^n}, cl(\bullet)$ – замыкание

$B_n = \bigsqcup\limits_{k=1}^{k_n} (a_{nk} + \delta, b_{nk}) \subset A_n$

$cl(B_n) = \bigsqcup\limits_{k=1}^{k_n} [a_{nk} + \delta, b_{nk}) \subset A_n$

$P(A_n) - P(B_n) = \sum\limits_{k=1}^{k_n} F(b_{nk}) - F(a_{nk} - F(b_{nk}) + F(a_{nk} + \delta)) = \sum\limits_{k=1}^{k_n} (F(a_{nk} + \delta) - F(a_{nk})) < \dfrac{\varepsilon}{2^n}$ – при правильном δ, k_n
$cl(B_n) \subset A_n$
$\bigcap cl(B_n) \subset \varnothing$
$\bigcap cl(B_n) = \varnothing$
$\bigcup\limits_{n=1}^{\infty} \overline{cl(B_n)} = \mathbb{R} \Rightarrow \bigcup\limits_{n=1}^{\infty} \overline{cl(B_n)} \supset [-M, M]$
$cl(B_n)$ – открыто
$[-M, M]$ – компакт
Тогда $\bigcup\limits_{n=1}^{N} \overline{cl(B_n)} \supset [-M, M]$
Тогда $\bigcap\limits_{n=1}^{N} cl(B_n) = \varnothing$

$\bigcap\limits_{n=1}^{N} B_n = \varnothing$

$P(A_n) = P(A_N \setminus \bigcap\limits_{n=1}^{N} B_n) + P(\bigcap\limits_{n=1}^{N} B_n) = P(\bigcup\limits_{n=1}^{N} (A_N \setminus B_n)) \le \sum\limits_{n=1}^{N} P(A_N \setminus B_n) \le$

$\sum\limits_{n=1}^{N} P(A_n \setminus B_n) = \sum\limits_{n=1}^{N} P(A_n) - P(B_n) \le \sum\limits_{n=1}^{N} \frac{\varepsilon}{2^n} < \varepsilon$
Тогда $P(A_n) \to 0$
Теперь пусть A_n – не ограниченные
$\varepsilon > 0$
$\exists\,[-M, M] : P(X \in \overline{[-M, M]}) < \varepsilon$
$N : F(N) > 1 - \frac{\varepsilon}{2}, F(-N) < \frac{\varepsilon}{2}$
$F(N) = F(-N) + P((-N, N]) > 1 - \frac{\varepsilon}{2}$
$F(N) = 1 - P((N, +\infty))$
$P((N, +\infty)) < \frac{\varepsilon}{2}$
$P(A_n) = P(A_n \cap [-M, M]) + P(A_n \cap \overline{[-M, M]}) \le P(\overline{[-M, M]}) < \varepsilon$
Т.о. P – вероятность
$X(\omega) := \omega$
$F_X(t) = P((-\infty, t]) = F(t)$
Замечание

5 Дискретные случайные величины и распределения

Определение
X, P_x – дискретные, если существует не более чем счетное $E : P(X \in E) = 1$
$F(t) = P(X \le t) = \sum\limits_{k} p(k)\mathbb{1}(t \ge x_k), \mathbb{1}(cond) = (int)(cond)$
Примеры

1. Вырожденное:
   $P(X = c) = 1$
   $X \sim I(c), I_c$

2. Дискретное равномерное на $\{x_1, \ldots, x_n\}$
   $P(X = x_k) = \frac{1}{n}$
   $DU(x_1, \ldots, x_n)$

3. Распределение Бернулли
   $Bern(p), p \in (0,1)$
   $P(X=0) = q = 1-p, P(X=1) = p$

4. Биномиальное; $Bin(n,p)$
   $P(X=k) = \binom{n}{k} p^k q^{n-k}, k \in \{0, \dots, n\}$

5. Полиномиальное; $Poly(n,p)$
   $p = (p_1, \dots, p_m)$ – вектор вероятностей
   $P(S_1 = k_1, \dots, S_m = k_m) = \binom{n}{k_1, \dots, k_m} p_1^{k_1} \dots p_m^{k_m}$

6. Геометрическое; $Geom(p)$
   X – количество неудач до первого успеха
   $P(X=k) = q^k p, k \in \mathbb{N}_0$
   Альтернативная интерпретация геометрического распределения
   X – номер первого успеха
   $P(X=k) = q^{k-1} p, k \in \mathbb{N}$

7. Отрицательное биномиальное; $NB(r,p), p \in (0,1), r > 0$

   - $r \in \mathbb{N}:$
     $X \sim NB(r,p) \Leftrightarrow X$ – номер r-ого успеха (начало отсчета в r)
     $P(X=k) = P(\text{успех с номером } r \text{ случился на шаге } k+r) = \binom{k+r-1}{r-1} p^r q^k$
   - Если $r \in \mathbb{R}$, то $P(X=k) = \dfrac{\Gamma(k+1)}{\Gamma(r)k!} p^r q^k$

8. Распределение Пуассона: $Pois(\lambda), \lambda > 0$
   $P(X=k) = e^{-\lambda} \dfrac{\lambda^k}{k!}, k \in \mathbb{N}_0$

9. Гипергеометрическое
   $HG(M, N, K)$
   $M \in \mathbb{N}$ – количество деталей
   $N \in [1:M]$ – количество «хороших» деталей
   $K \in [1:M]$ – количество деталей, которые мы вытаскиваем (без

возвращения)
X – количество «хороших» деталей, которые мы вытащили
$P(X=j)=\frac{\binom{N}{j}\binom{M-N}{K-j}}{\binom{M}{K}}, \max(K+N-M,0)\leq j\leq \min(N,K)$
Пусть $M\to\infty, \frac{N}{M}\to p$
$P(X=j)=\frac{N!(M-N)!K!(M-K)!}{j!(N-j)!(K-j)!(M-N-K+j)!M!}$
$=\binom{K}{j}\frac{N(N-1)\dots(N-j+1)}{M(M-1)\dots(M-N+1)}\frac{(M-K)\dots(M-N-K+j+1)}{}\to$
$\binom{K}{j}p^j(1-p)^{k-j}$

$EX=\sum\limits_{j=\max(0,N+K-M)}^{\min(N,K)} j\frac{\binom{N}{j}\binom{M-N}{K-j}}{\binom{M}{K}}=$
$\sum\limits_{j=\max(1,N+K-M)}^{\min(N,K)} j\frac{\binom{N}{j}\binom{M-N}{K-j}}{\binom{M}{K}}=$
$N\sum\limits_j\frac{(N-1)!}{(j-1)!(N-j)!}\frac{\binom{M-N}{K-j}}{\binom{M}{K}}=$
$\frac{NK}{M}\sum\limits_j\frac{\binom{N-1}{j-1}\binom{M-N}{K-j}}{\binom{M-1}{K-1}}=$
$\frac{NK}{M}\underbrace{\sum\limits_{i=\max(0,N+K-M-1)}^{\min(N-1,K-1)}\frac{\binom{N-1}{i}\binom{M-N}{K-i-1}}{\binom{M-1}{K-1}}}_{1}=\frac{NK}{M}$

6 Абсолютно непрерывные распределения

Определение
X, P_X – абсолютно непрерывные, если $\exists\, p(x):\mathbb{R}\to\mathbb{R}, p(x)\geq 0, p(x)$ – интегрируемо по Лебегу на $\mathbb{R}$
$P(X\in B)=\int_B p(x)\,\mathrm{d}\,x, B\in\mathcal{B}$
p – плотность
Теорема

1. $P(X=c)=0$, т.к. $P(X=c)=\int_{\varnothing}\ldots=0$

2. (a) $F(x)=P(X\le x)=\int_{-\infty}^{x} p(t)\,\mathrm{d}\,t$

   (b) $F(x)$ – непрерывна (равномерно непрерывная)

   (c) $F'(x)=p(x)$ почти везде

3. $\int_{-\infty}^{\infty} p(x)\,\mathrm{d}\,x=1$

4. $P(X\in(x_0,x_0+h))=F(x_0+h)-F(x_0)=p(x_0)h+o(h)\approx p(x_0)h$ при малых h

Определение (носитель)
E – носитель $(\operatorname{supp}P_X)\Leftrightarrow P(X\in E)=1, E=\overline{E}, E$ наименьшее по включению
Замечание
λ – мера Лебега
ν – другая мера, заданная на $\mathcal{B}, \nu(\mathbb{R})<+\infty$
ν – абсолютно непрерывна относительно $\lambda\Leftrightarrow\lambda(A)=0\Rightarrow\nu(A)=0$
$\nu\ll\lambda$
Теорема Радона-Никодима
$\nu\ll\lambda\Rightarrow\exists\, f(x):\nu(B)=\int_B f(x)\lambda(\mathrm{d}\,x)$
Пример

1. $U[a,b]$: Равномерное на $[a,b]$
   $p(x)=\dfrac{1}{b-a}\mathbb{1}(x\in[a,b])$
   $F(t)=\int_{-\infty}^{t} p(x)\,\mathrm{d}\,x=\begin{cases} 0, & t<a\\ \dfrac{t-a}{b-a}, & a\le t\le b\\ 1, & t>b\end{cases}$
   $X\sim U[a,b]$
   $Y=cX+d, c>0$
   $Y\sim U[ca+d, cb+d]$
   $F_Y(t)=P(Y\le t)=P(cX+d\le t)=P(X\le\dfrac{t-d}{c})=F_X(\dfrac{t-d}{c})$

2. $N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma > 0$: Нормальное распределение
$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2})$
μ – высота графика, σ – ширина
$N(0,1)$ – стандартный нормальный закон
$\phi = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}), \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}$
$\Phi(-x) = 1 - \Phi(x)$
Доказательство
$\Phi(x) = \int_{-\infty}^{-x} \ldots + \int_{-x}^{x} \ldots = \int_{-\infty}^{-x} \ldots + 2\int_{0}^{x} \ldots = \int_{-\infty}^{-x} \ldots + 2(\int_{-\infty}^{x} \ldots - \int_{-\infty}^{0} \ldots) = \Phi(-x) + 2(\Phi(x) - \frac{1}{2})$
$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} \,\mathrm{d}\, t$ – функция Лапласа; $\Phi_0(-x) = -\Phi_0(x)$
$\Phi(x) = \frac{1}{2} + \Phi_0(x)$
$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} \,\mathrm{d}\, t$
$\Phi_0(x) = \frac{1}{2} \mathrm{erf}(\frac{x}{\sqrt{2}})$
$y_{\text{изм}} = y_{\text{реал}} + \varepsilon, \varepsilon \sim N(0, \sigma^2)$
$P(|y_{\text{изм}} - y_{\text{реал}}| > \delta) = P(|\varepsilon| > \delta)$
$X \sim N(\mu, \sigma^2)$
$Y = aX + b$
$F_Y(t) = P(aX + b \leq t) \underset{a>0}{=} P(X \leq \frac{t-b}{a}) = F_x(\frac{t-b}{a})$
$p_Y(t) = (F_X(\frac{t-b}{a}))'_t = p_X(\frac{t-b}{a})\frac{1}{a} = \ldots$
$Y \sim N(a\mu + b, \sigma^2 a^2)$
Для $a < 0$ аналогично
$X \sim N(\mu, \sigma^2), Y = aX + b \Rightarrow Y \sim N(a\mu + b, a^2\sigma^2)$
$U = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} - \frac{\mu}{\sigma} \Rightarrow U \sim N(0,1)$

$X \sim N(\mu, \sigma^2)$
$P(a \leq X \leq b) = P(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$
$P(|X - \mu| < k\sigma) = P(-k\sigma < X - \mu < k\sigma) = P(\frac{X-\mu}{\sigma} \in (-k, k)) =$

$\Phi(k) - \Phi(-k) = 2\Phi(k) - 1$

3. $\mathrm{Exp}(\lambda), \lambda > 0$: Экспоненциальнле/показательное
   $p(x) = \lambda e^{-\lambda x}\mathbb{1}(x \geq 0)$
   $F(x) = (1 - e^{-\lambda x})\mathbb{1}(x \geq 0)$

4. $\Gamma(\alpha, \beta), \alpha, \beta > 0$: Гамма-распределение
   $p(x) = \dfrac{\beta^\alpha x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)}\mathbb{1}(x \geq 0)$
   $\Gamma(1, \beta) = \mathrm{Exp}(\beta)$
   $\beta - rate$
   $\alpha \in \mathbb{N} \Rightarrow$ распределение Эрланга порядка k

5. $\mathrm{Cauchy}(x_0, \gamma)$: Распределение Коши
   $x_0 \in \mathbb{R}$ – сдвиг
   $\gamma > 0$ – scale
   $p(x) = \dfrac{1}{\pi\gamma}\dfrac{1}{1 + (\frac{x-x_0}{\gamma}^2)}$
   $F_X(t) = \dfrac{1}{\pi\gamma}\displaystyle\int_{-\infty}^{t}\dfrac{\mathrm{d}\,x}{1 + (\frac{x-x_0}{\gamma})^2} = \left|\begin{array}{c}\dfrac{x-x_0}{\gamma} = a \\ \mathrm{d}\,x = \gamma\,\mathrm{d}\,u\end{array}\right| = \dfrac{1}{\pi}\displaystyle\int_{-\infty}^{\frac{t-x_0}{\gamma}}\dfrac{\mathrm{d}\,u}{1+u^2} =$
   $\dfrac{1}{\pi}\operatorname{arctg} u\Big|_{-\infty}^{\frac{t-x_0}{\gamma}} = \dfrac{1}{\pi}\operatorname{arctg}\dfrac{t-x_0}{\gamma} + \dfrac{1}{2}$
   Пример
   $X \sim \mathrm{Cauchy}(0, 1)$
   $Y = \dfrac{1}{X}$
   $F_Y(t) = P(\dfrac{1}{X} \leq t) = P(\dfrac{1}{X} \leq t, X < 0) + P(\dfrac{1}{X} \leq t, X > 0)$
   $P(\dfrac{1}{X} \leq t, X < 0):$
   $t \geq 0 \Rightarrow P(\ldots) = P(X < 0) = \dfrac{1}{2}$
   $t < 0 \Rightarrow P(X \leq 1, X < 0) = P(X \geq \dfrac{1}{t}, X < 0) = F(0) - F(\dfrac{1}{t}) =$
   $\dfrac{1}{2} - \dfrac{1}{2} - \dfrac{1}{\pi}\operatorname{arctg}\dfrac{1}{t} = -\dfrac{1}{\pi}\operatorname{arctg}\dfrac{1}{t}$
   $P(\dfrac{1}{X} \leq t, X > 0):$
   $t \leq 0 \Rightarrow P(\ldots) = 0$

$$t > 0 \Rightarrow P(X \geq \frac{1}{t}, X > 0) = P(X \geq \frac{1}{t}) = 1 - F(\frac{1}{t}) = \frac{1}{2} - \frac{1}{\pi}\operatorname{arctg}\frac{1}{t}$$

Т.о. $P(\frac{1}{X} \leq t, X > 0) = \begin{cases} -\frac{1}{\pi}\operatorname{arctg}\frac{1}{t}, & t < 0 \\ \frac{1}{2}, & t = 0 \\ 1 - \frac{1}{\pi}\operatorname{arctg}\frac{1}{t}, & t > 0 \end{cases}$

Теорема
Пусть p_x – плотность с.в. X
$Y = g(X), g \in C^1$
g' строго монотонна
Тогда $p_Y(y) = p_x(g^{-1}(y))|\frac{\partial}{\partial y}(g^{-1}(y))| = p_x(g^{-1}(y))\frac{1}{|g'_x(g^{-1}(y))}$
Доказательство
Пусть $g' \geq 0$
$F_Y(t) = P(g(X) \leq t) = P(X \leq g^{-1}(t)) = F_X(g^{-1}(t))$
$p_Y(t) = (F_Y(t))'_t = p_X(y^{-1}(t))\frac{\partial g^{-1}(t)}{\partial y}$
Пусть $g' < 0$
$F_Y(t) = 1 - F_X(g^{-1}(t))$
$p_Y(t) = -p_X(y^{-1}(t))\frac{\partial g^{-1}(t)}{\partial y}$

7 Сингулярное распределение

Определение
X, P_X – сингулярное $\Leftrightarrow \operatorname{supp} P_X = E$ & $\lambda E = 0$ и $\forall x \in E$ $P(X \in E) = 0$
Замечание
Сингулярное $\Leftrightarrow \lambda\{x : F(x+\varepsilon) - F(x-\varepsilon) > 0\} = 0\ \forall \varepsilon$
Теорема Лебега о разложении F
F – функция распределения
Тогда $\exists c_1, c_2, c_3 \geq 0 : c_1 + c_2 + c_3 = 1, F(x) = c_1F_1(x) + c_2F_2(x) + c_3F_3(x), F_1$ – функция распределения дискретной величины, F_2 – функция распределения, абсолютно непрерывная, F_3 – функция распределения сингулярного закона
Замечание
У F могут быть только скачки, не более чем счетное количество

8 Случайные векторы

Определение
$X = \{X_1, \dots, X_n\}$ – случайный вектор (многомерная случайная величина), если X_i – случайная величина
$P_X(X_1 \in B_1, \dots, X_n \in B_n)$ – распределение случайного вектора (совместное распределение случайных величин $X_1, \dots, X_n, B_i \in \mathcal{B}$)
$P_X : \sigma(\mathcal{B}^n) \to \mathbb{R}$
Определение
$F(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$ – совместная функция распределения
Свойства

1. $x_i \to -\infty$ для некоторого $i \Rightarrow F(x_1, \dots, x_n) \to 0$
2. $x_i \to +\infty$ для всех $i \Rightarrow F(x_1, \dots, x_n) \to 1$
3. F – непрерывна справа по каждой переменной
4. $\Delta_i F(x_1, \dots, x_i + \delta_i, \dots, x_n) - F(x_1, \dots, x_n)$
   $\Delta_1, \dots, \Delta_n F \geq 0 \; \forall \, x_1, \dots, x_n \; \forall \, \delta_1, \dots \delta_n \geq 0$

Теорема
F – функция, для которой справедливы свойства $1 - 3$
Тогда существует вероятностное пространство и случайная величина с данной функцией распределения
Определение
X_i, P_X – дискретно, если $\operatorname{supp} P_X$ – не более чем счетный
Замечание
$\operatorname{supp} P_X \neq \times_{i=1}^n \operatorname{supp} P_x$
Одномерное распределение = *Маргинальное* распределение
Пример
Полиномиальное: $Poly(n, p)$
$n \in \mathbb{N}$
$p = (p_1, \dots, p_n), \sum p_j = 1$
$S_n = (S_{n,1}, \dots, S_{n,m})$
$P(S_{n,1} = k_1, \dots, S_{nm} = k_m) = \dfrac{n!}{k_1! \dots k_m!} p_1^{k_1} \dots p_m^{k_m}$
$\sum k_i = n, k_i \geq 0$

$S_{nj} \sim Bin(n, p_j)$

Определение

X, P_X – абсолютно непрерывная $\Leftrightarrow \exists\, p(x_1, \dots, x_n) \geq 0$ – плотность распределения

$P(X \in B) = \displaystyle\int_B p(x)\,\mathrm{d}\,x$

$F(x_1, \dots, x_n) = \displaystyle\int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} p(t_1, \dots, t_n)\,\mathrm{d}\,t_1 \dots \mathrm{d}\,t_n$

$p(t_1, \dots, t_n) = \dfrac{\partial^n F}{\partial x_1 \dots \partial x_n}$

$P(G(X) \in B) = \displaystyle\int_{x: F(x) \in B} p(x)\,\mathrm{d}\,x, G$ – некая функция

Определение (равномерное распределение на подмножестве)

$U(E) : p(x) = \dfrac{1}{\mu E}\mathbb{1}(x \in E), \mu$ – мера

Теорема (о плотности функции от непрерывного случайного вектора)

Пусть $g : \mathbb{R}^n \to \mathbb{R}^n, g$ – гладкая биекция

Тогда $p_{g(U)}(t) = p_U(g^{-1}(t))|\det(g^{-1})'(t)| = \dfrac{p_U(g^{-1}(t))}{|\det g'(g^{-1}(t))|}$

Определение (многомерное нормальное распределение/гауссовский вектор)

1. $X = (X_1, X_n)$ – стандартный гауссовский вектор, если $p_X = \dfrac{1}{(\sqrt{2\pi})^n}\exp(-\dfrac{1}{2}\|x\|^2), \|x\|^2 = x^T x = \sum x_i^2$
   $X \sim N(0, E_n), 0 \in \mathbb{R}^n$
2. $X \sim N(0, E_n), Y = AX + b, A \in M_{m \times n}, b \in \mathbb{R}^m \Rightarrow Y \sim N(b, AA^T)$
3. $U \sim N(\mu, \Sigma)$, где $\mu \in \mathbb{R}^n, \Sigma \in M_{n \times n}$
   $\Sigma = \Sigma^T, \Sigma > 0$
   $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n), \lambda_i \geq 0$
   $\Lambda = L^T \Sigma L, L = (u_1, \dots, u_n)$ – ОНБ, $L^T L = E$
   $\Sigma = L\Lambda L^T = L = \underbrace{L\sqrt{\Lambda}L^T}_{\sqrt{\Sigma}}\underbrace{L\sqrt{\Lambda}L^T}_{\sqrt{\Sigma}}$
   $U = \sqrt{\Sigma}X + \mu, X \sim N(0, E)$
   $\Sigma > 0 \Rightarrow p(x_1, \dots, x_n) = \dfrac{1}{(\sqrt{2\pi})^n\sqrt{\det\Sigma}}\exp -\dfrac{1}{2}(x - \mu)^T\Sigma^{-1}(x - \mu)$

$X = \sqrt{\Sigma}^{-1}(X-\mu)$
$Y' = (\sqrt{\Sigma})^{-1}$
$\det\Sigma = \det\sqrt{\Sigma}^2$
$|\det\sqrt{\Sigma}| = \sqrt{\det\Sigma}$
$p_U(t) = \dfrac{p_X(g^{-1}(t))}{|\det g'(g^{-1}(t))|} = \dfrac{1}{(\sqrt{2\pi})^n}\exp(-\dfrac{1}{2}(\sqrt{\Sigma})^{-1}(t-\mu)^T(\sqrt{\Sigma})^{-1}(t-\mu))\dfrac{1}{|\det\sqrt{\Sigma}|}$
$= \dfrac{1}{\sqrt{(2\pi)^n\det\Sigma}}\exp(-\dfrac{1}{2}(t-\mu)^T\Sigma^{-1}(t-\mu))$

9 Независимые случайные величины

Определение
$X_1, \dots, X_n$ – независимые, если $\forall\, B_1, \dots, B_n$ $P(X_1 \subset B_1, \dots, X_n \subset B_n) = P(X_1 \subset B_1) \dots P(X_n \subset B_n)$
Последовательность независима, если любой ее префикс независимый
Теорема (общий критерий независимости)
Величины $X_1, \dots, X_n$ независимы $\Leftrightarrow F(x_1, \dots, x_n) = F_1(x_1)\dots F_n(x_n)$
Доказательство $\Rightarrow$ очевидно
Доказательство $\Leftarrow$
Докажем для размерности 2
$P(X_1 \in (a_1, a_1+\delta_1], X_2 \in (a2, a_2+\delta_2]) =$
$\Delta_1\Delta_2 F(a_1, a_2) = \Delta_1[F(a_1, a_2+\delta_2) - F(a_1, a_2)] =$
$F(a_1+\delta_1, a_2+\delta_2) - F(a_1, a_2+\delta_2) - F(a_1+\delta_1, a_2) + F(a_1, a_2) =$
$F_1(a_1+\delta_1)F_2(a_2+\delta_2) - F_1(a_1)F_2(a_2+\delta_2) - F_1(a_1+\delta_1)F_2(a_2) + F_1(a_1)F_2(a_2) =$
$(F_1(a_1+\delta_1) - F_1(a_1))(F_2(a_2+\delta_2) - F_2(a_2)) =$
$P(X_1 \in (a_1, a_1+\delta_1])P(X_2 \in (a_2, a_2+\delta_2])$
Теорема (критерий независимости дискретных величин)
Дискретные $X_1, \dots, X_n$ – независимые $\Leftrightarrow P(X_1 \in \{x_{1,1}, \dots, x_{1,m_1}\}, \dots, X_n \in \{x_{n,1}, \dots, x_{n,m_n}\}) = \prod\limits_{j=1}^{n} P(X_j \in \{x_{j,1}, \dots, x_{j,im_j}\})$
Доказательство $\Rightarrow$ очевидно
Доказательство $\Leftarrow$
$P(X_1 \in B_1, \dots, X_n \in B_n) = \sum\limits_{i_1,\dots,i_n} P(X_1 = x_{1,i_1}, \dots, X_n = x_{n,i_n}) = \sum\limits_j \prod\limits_i P(X_j =$

$x_{j,i}) = \prod_i \sum_j P(X_j = x_{j,i}) = \prod_i P(X_i \in B_i)$

Теорема (критерий независимости непрерывных величин)

Непрерывные $X_1, \ldots, X_n$ – независимые $\Leftrightarrow p(x_1, \ldots, x_n) = \prod_i p_i(x_i)$

Доказательство $\Rightarrow$

$F(x_1, x_2) = F_1(x_1)F_2(x_2) = (\int_{-\infty}^{x_1} p_1(x)\,\mathrm{d}\,x)(\int_{-\infty}^{x_2} p_2(y)\,\mathrm{d}\,y) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} p_1(x)p_2(y)\,\mathrm{d}\,x\,\mathrm{d}\,y$

Доказательство $\Leftarrow$ очевидно (наверное)

Пример

$\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2), \sigma_i > 0$

$p(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \sigma_1^2 \ldots \sigma_n^2}} \exp(-\frac{1}{2} x^T (\sigma_1^2 \ldots \sigma_n^2) x) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp(-\frac{1}{2}\frac{x_i^2}{\sigma_i^2})$

Лемма

X, Y – независимые целочисленные

Тогда $P(X + Y = K) = \sum_i P(X = i)P(Y = K - i) = \sum_i P(X = K - i)P(Y = i)$

Доказательство

$P(X + Y = K) = \sum_i P(X + Y = K | X = i)P(X = i) = \sum_i P(X = K - i)P(X = i)$

Пример

1. $S_n \sim Bin(n, p)$
   $X_n = X_1 + \ldots + X_n, X_i = Bern(p), X_i$ – независимые
   $P(S_{n+1} = k) = P(S_n = k)P(X_{n+1} = 0) + P(S_n = k - 1)P(X_{n+1} = 1) = \binom{n}{k} p^k q^{n-k+1} + \binom{n}{k-1} p^k q^{n-k+1} = \binom{n+1}{k} p^k q^{n-k+1}$

2. $X \sim Pois(\lambda)$
   $Y = Pois(\mu)$
   $X + Y \sim Pois(\lambda + \mu)$

3. $X \sim NB(r_1, p), Y \sim NB(r_2, p)$
   Тогда $X + Y \sim NB(r_1 + r_2, p)$
   $X_1 + \ldots + X_n \sim NB(n, p), X_i \sim NB(1, p) = Geom(p)$
   $P(X_1 + X_2 = K) = \binom{(K+2) - 1}{1} q^k p^2$

$$P(S_n + X_{n+1} = K) = \binom{n+k}{n} q^k p^n$$

Лемма

X, Y – независимые

Тогда $p_{X+Y}(t) = \int_{\mathbb{R}} p_X(u) p_Y(t-u) \,\mathrm{d}\, u$

Доказательство

$F_{X+Y}(t) = P(X+Y \le t) = \int_{-\infty}^{\infty} p_X(u) \int_{-\infty}^{t-u} p_Y(v) \,\mathrm{d}\, v$

Пример

1. X, Y – независимые, $X \sim N(\mu_X, \sigma_X^2), Y \sim N(\mu_Y, \sigma_Y^2)$
   Тогда $X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$
   $X + Y = \sigma_X U_X + \mu_x + \sigma_Y U_Y + \mu_Y, U_X, U_Y \sim N(0,1)$
   $\sigma_Y^2 u^2 + \sigma_X^2 (t-u)^2 = \underbrace{(\sqrt{\sigma_X^2 + \sigma_Y^2} u - \frac{\sigma_X^2 t}{\sqrt{\sigma_X^2 + \sigma_Y^2}}}_{y})^2 + \frac{t^2 \sigma_X^2 \sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}$
   $P_{\sigma_X U_X + \mu_x + \sigma_Y U_Y + \mu_Y}(t) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \frac{1}{\sqrt{2\pi\sigma_Y^2}} \int_{-\infty}^{\infty} \exp(-\frac{1}{2}(\frac{u^2}{\sigma_X^2} + \frac{(t-u)^2}{\sigma_Y^2})) \,\mathrm{d}\, u =$
   $= \frac{1}{\sqrt{2\pi\sigma_X^2}} \frac{1}{\sqrt{2\pi\sigma_Y^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \exp(-\frac{1}{2(\sigma_X^2 + \sigma_Y^2)}(y^2 + \frac{t^2 \sigma_X^2 \sigma_Y^2}{\sigma_X^2 + \sigma_Y^2})) \,\mathrm{d}\, u =$
   $= \frac{1}{\sqrt{2\pi(\sigma_X^2 + \sigma_Y^2)}} \exp(-\frac{1}{2} \frac{t^2}{\sigma_X^2 + \sigma_Y^2}) \underbrace{(\frac{1}{2\pi\sigma_X^2\sigma_Y^2} \int_{-\infty}^{\infty} \exp(-\frac{y^2}{2\sigma_X^2\sigma_Y^2}) \,\mathrm{d}\, y)}_{1}$
   $p_{U+\mu_x+\mu_Y}(t) = p_U(t - \mu_x - \mu_y)|t'| = \frac{1}{\sqrt{2\pi\sigma_X^2\sigma_Y^2}} \exp(-\frac{1}{2} \frac{(t - \mu_X - \mu_Y)^2}{\sigma_X^2 + \sigma_Y^2})$

Утверждение

$X_1, \ldots, X_n$ – независимые

$X_i \sim Exp(\lambda)$

Тогда $X_1 + \ldots + X_n \sim \Gamma(n, \lambda)$ – распределение Эранга порядка n

$P_{X_1+X_2}(t) = \int_{-\infty}^{+\infty} p_1(x) p_2(t-x) \,\mathrm{d}\, x = \lambda^2 t e^{-\lambda t} \mathbb{1}(t \ge 0)$

$P_{S_n + X_{n+1}} = \int_{-\infty}^{+\infty} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \mathbb{1}(x \ge 0) \lambda e^{-\lambda(t-x)} \mathbb{1}(t - x \ge 0) \,\mathrm{d}\, x = \frac{\lambda^{n+1} x^n e^{-\lambda t}}{n!} \mathbb{1}(t \ge 0)$

10 Симуляция (моделирование случайных величин)

Пусть есть $rand() : U[0, 1]$ – независимые

```
def Bern(p):
    return rand() <= p


def Bin(n, p):
    return sum(Bern(p) for _ in range(n))

def Geom(p):
    res = 0
    while(!Bern(p)):
        res += 1
    return res

def Discrete(p_1, p_2, ..., p_n, x_1, ..., x_n):
    v = rand()
    if p_1 + ... + p_{n-1} < v: return x_n
    if p_1 + ... + p_{n-2} < v: return x_{n-1}
    ...
    if p_1 < v: return x_2
    return x_1

def Poly(1, p): ... #аналогично

def Poly(n, p):
    return sum(Poly(1, p) for _ in range(n))

def U(a, b):
    return a + (b - a) * rand()

def Exp(lmbda):
    return -ln(rand())/lmbda

def Gamma(n, lmbda):
    return sum(Exp(lmbda) for _ in range(n))

# Случайная величина с распределением F
def Generic(F):
    G = reversed(F) #обратная функция
    return G(rand())
```

```python
    def N(0, 1):
        # n -- какоето- большое, p -- любое
        return (Bin(n, p) - n*p) / sqrt(n * p * (1 - p))
```

Утверждение
$X_1, \ldots, X_n \sim Exp(1)$ – независимые
$X = \max\{n : X_1 + \ldots + X_n < \lambda\}, \lambda \geq 0$
Тогда $X \sim Pois(\lambda)$
Доказательство

$$P(X \leq k) = P(\underbrace{X_1 + \ldots + X_{k+1}}_{\sim\Gamma(k+1,1)} \geq \lambda) = \int_\lambda^{+\infty} \frac{x^k e^{-x}}{k!} \,\mathrm{d}\, x = -\frac{x^k e^{-x}}{k!}\Bigg|_\lambda^{+\infty} + \int_\lambda^{+\infty} x^{k+1} e^{-x} (k-1)! \,\mathrm{d}\, x = \frac{\lambda^k e^{-\lambda}}{k!} + \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!} + \ldots + \frac{\lambda^0 e^{-\lambda}}{0!}$$

11 Вероятностные интегралы

$(\Omega, \mathcal{A}, P) \overset{X}{\mapsto} (\mathbb{R}^n, \mathcal{B}, P_X)$
$P_X(B) = P(\{\omega : X(\omega) \in B\})$
$g \geq 0$ – *борелевская*, т.е. $\forall\, B \in \mathcal{B}\ g^{-1}(B) \in \mathcal{B}$

$$\underbrace{\int_\Omega g(X(\omega)) P(\mathrm{d}\,\omega)}_{\lim_{\operatorname{diam} V(\omega_k)\to 0} \sum g(X(\omega_k)) P(V(\omega_k))} = \underbrace{\int_{\mathbb{R}^n} g(x) P_X(\mathrm{d}\, x)}_{\lim_{\operatorname{diam} V(x_k)\to 0} \sum g(x_k) P(X \in V(x_k))}$$

$, V(\cdot)$ – какая-то окрестность

Интеграл Стилтьеса

$$\int_{\mathbb{R}} g(x) \,\mathrm{d}\, F(x) = \lim_{x_{k+1}-x_k \to 0, x_k^* \in [x_k, x_{k+1}]} \sum g(x_k^*)(F(x_{k+1}) - F(x_k)) =$$

$$= \begin{cases} \sum g(x_i) P(X = x_i), & \text{дискретный случай} \\ \int g(x) p(x) \,\mathrm{d}\, x, & \text{непрерывный случай} \end{cases}$$

$g \geq 0 \Rightarrow$ интеграл либо конечный, либо $+\infty$

$$g = g_+ - g_i, I = \int_{\mathbb{R}^n} g_+(x) P(\mathrm{d}\, x) - \int_{\mathbb{R}^n} g_-(x) P(\mathrm{d}\, x)$$

- $I_+, I_- < +\infty \Rightarrow I = I_+ - I_-$
- $I_+ = +\infty, I_i < +\infty \Rightarrow I = +\infty$

- $I_+ < +\infty, I_i = +\infty \Rightarrow I = -\infty$
- $I_+, I_i = +\infty \Rightarrow I$ не определено

Свойства

- $\displaystyle\int af + bg = a\int f + b\int g$
- $\displaystyle\int_{A\sqcup B} = \int_A + \int_B$
- $\displaystyle f \le g \Rightarrow \int f \le \int g$

Теорема Фубини для независимых величин

X, Y– независимые $\Rightarrow g(x,y)P(\mathrm{d}\,x, \mathrm{d}\,y) = \displaystyle\int_{\mathbb{R}}(\int_{\mathbb{R}} g(x,y)P_Y(\mathrm{d}\,y))P_X(\mathrm{d}\,x)$

Пример (свертка распределений)

X, Y – независимые

$F_{X+Y}(t) = P(X+Y \le t) = \displaystyle\int_{\mathbb{R}^2} \mathbb{1}(x+y \le t)P(\mathrm{d}\,x, \mathrm{d}\,y) = \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}(x+y \le t)P_X(\mathrm{d}\,x)P_Y(\mathrm{d}\,y) = \int_{\mathbb{R}} P(X \le t-y)P_Y(\mathrm{d}\,y) = \int_{\mathbb{R}} F_X(t-y)\,\mathrm{d}\,F_Y(y)$

Аналогично $F_{X+Y}(t) = \displaystyle\int_{\mathbb{R}} F_Y(t-x)\,\mathrm{d}\,F_X(x)$

12 Математическое ожидание

Определение

$EX = \displaystyle\int_{\mathbb{R}} x\,\mathrm{d}\,F(x) = \int xP_X(\mathrm{d}\,x)$

Дискретный случай: $EX = \sum x_k p_k$

Непрерывный случай: $EX = \displaystyle\int xp(x)\,\mathrm{d}\,x$

Свойства

1. $Ef(X_1, \ldots, X_n) = \displaystyle\int_{\mathbb{R}^n} f(x_1, \ldots, x_n)P(\mathrm{d}\,x_1, \ldots, \mathrm{d}\,x_n) = \int_{\mathbb{R}} y\,\mathrm{d}\,F_Y(y), y = f(x_1, \ldots, x_n)$
2. $E(aX) = aEx$

3. $E(X+Y) = EX + EY$
4. X, Y – независимые $\Rightarrow E(XY) = (EX)(EY)$
5. $\underbrace{X \geq 0}_{P(X \geq 0)=1} \Rightarrow EX \geq 0$
6. $\underbrace{X \geq Y}_{P(X \geq Y)=1} \Rightarrow EX \geq EY$
7. **Теорема (неравенство Маркова)**
   $X \geq 0 \Rightarrow P(X \geq C) \leq \dfrac{EX}{c}$
   Доказательство
   $EX = \displaystyle\int_0^{+\infty} d \,\mathrm{d}\, F(x) = \int_0^c x \,\mathrm{d}\, F(x) + \int_{c+\infty} x \,\mathrm{d}\, F(x) \geq 0 + c\int_c^{+\infty} \mathrm{d}\, F(x) =$
   $cP(X \geq c)$
8. $X \geq 0, EX = 0 \Rightarrow \underbrace{X = 0}_{P(X=0)=1}$
   Доказательство
   $P(X \geq c) \leq \dfrac{0}{c} = 0 \Rightarrow P(X < c) = 1 \forall\, c \Rightarrow P(X = 0) = 1$
9. $\operatorname{supp} X = \mathbb{N} \Rightarrow EX = \displaystyle\sum_{k=1}^{+\infty} P(X \geq k)$
10. $\operatorname{supp} X = [0, +\infty)$
    $EX = \displaystyle\int_0^{+\infty} P(X \geq r) \,\mathrm{d}\, r$
    Доказательство
    $\displaystyle\int_0^{+\infty} P(X \geq r) \,\mathrm{d}\, r = \int_0^{+\infty} (1 - F(r)) \,\mathrm{d}\, r = \underbrace{x(1 - F(x))\Big|_0^{+\infty}}_{0} + \int_0^{+\infty} xp(x) \,\mathrm{d}\, x$

Определение
X, Y– некоррелированные $\Leftrightarrow E(XY) = (EX)(EY)$
Определение

$$\operatorname{Var} X = E(X-EX)^2 = \begin{cases} \sum\limits_k (x_k - EX)^2 p_k \\ \int (x-EX)^2 p(x)\,\mathrm{d}\,x \end{cases}$$

$$Ef(x_1,\dots,x_n) = \int_{\mathbb{R}} f(x_1,\dots,x_n) P_X(\mathrm{d}\,x_1,\dots,\mathrm{d}\,x_n)$$

$\sqrt{\operatorname{Var} X}$ – среднее квадратическое(стандартное) отклонение

Свойства

1. $\operatorname{Var} X = EX^2 - (EX)^2$
   Доказательство
   $\operatorname{Var} X = E(X^2 - 2XEX + (EX)^2) = EX^2 - 2(EX)^2 + (EX)^2 = EX^2 - (EX)^2$

2. $\operatorname{Var} X \geq 0$

3. $\operatorname{Var} X = 0 \Rightarrow X = c$
   Доказательство
   $(X - EX)^2 = 0 \Rightarrow X = EX$

4. $\operatorname{Var}(aX+b) = a^2 \operatorname{Var} X$

5. $\operatorname{Var}(X \pm Y) = \operatorname{Var} X + \operatorname{Var} Y \pm 2(E(XY) - EX \cdot EY)$
   Доказательство
   $\operatorname{Var}(X \pm Y) = E(X^2 \pm 2XY + Y^2) - ((EX)^2 \pm 2EX \cdot EY + (EY)^2) = \operatorname{Var} X + \operatorname{Var} Y \pm 2(E(XY) - EX \cdot EY)$

6. $\operatorname{Var}(X_1 + \ldots + V_n) = \sum\limits_{i=1}^{n} \operatorname{Var} X_i + 2\sum\limits_{i<j}(E(X_iX_j) - EX_iEX_j)$

Теорема (неравенство Чебышева)

$P(|X-EX| \geq \varepsilon) \leq \dfrac{\operatorname{Var} X}{\varepsilon^2}$

Свойства

1. $P(|X-EX| < \varepsilon) \geq 1 - \dfrac{\operatorname{Var} X}{\varepsilon^2}$

2. $P(|X-EX| < k\sqrt{\operatorname{Var} X}) \geq 1 - \dfrac{1}{k^2}$

3. ЗБЧ (слабый)
   $X_1, \ldots, X_n$ – независимые
   $EX_i = \mathcal{M}$
   $\operatorname{Var} X_i = \sigma^2$
   $S_n = \sum\limits_{i=1}^{n} X_i$
   Тогда $P(|\frac{S_n}{n} - \mathcal{M}| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \to 0$
   $E\frac{S_n}{n} = \frac{1}{n}\sum EX_i = \mathcal{M}$
   $\operatorname{Var} \frac{S_n}{n} = \frac{n \operatorname{Var} X_1}{n^2} = \frac{\operatorname{Var} X_1}{n}$

Пример
Дискретные:

- $I(c)$
  $EX = 1c = c$
  $\operatorname{Var} X = 0$

- $Bern(p)$
  $EX = 1p + 0q = p$
  $EX^2 = p$
  $\operatorname{Var} X = p - p^2 = pq$

- $Bin(n, p)$
  $S_n = X_1 + \ldots + X_n$
  $X_i \sim Bern(p)$ – независимые
  $ES_n = nEX_1 = np$
  $\operatorname{Var} S_n = n \operatorname{Var} X_1 = npq$

- $Pois(\lambda)$
  $EX = \sum\limits_{k=1}^{\infty} ke^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\lambda\sum\limits_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda+\lambda}\lambda = \lambda$
  $E(X(X-1)) = EX^2 - EX = EX^2 - (EX)^2 + (EX)^2 - EX = \operatorname{Var} X + (EX)^2 - EX \Rightarrow \operatorname{Var} X = E(X(X-1)) - (EX)^2 + EX$
  $E(X(X-1)) = \sum\limits_{k=2}^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = \lambda^2$
  $\operatorname{Var} X = \lambda^2 - \lambda^2 + \lambda = \lambda$

- $Geom(p)$
  $EX = \sum\limits_{k=1}^{\infty} P(X \geq k) = \sum\limits_{k=1}^{\infty}\sum\limits_{j=1}^{k} P(X = j) = \sum\limits_{k=1}^{\infty}\sum\limits_{j=k}^{\infty} q^j p = \sum\limits_{k=1}^{\infty} \frac{q^k p}{1-q} = \sum\limits_{k=1}^{\infty} q^k = \frac{q}{p}$
  $E(X(X-1)) = \sum\limits_{k=2}^{\infty} k(k-1)q^k p = pq^2 \sum\limits_{k=2}^{\infty} k(k-1)q^{k-2} = pq^2(\sum\limits_{k=0}^{\infty} q^k)'' = pq^2(\frac{1}{1-q})'' = \frac{2pq^2}{(1-q)^3} = 2\frac{q^2}{p^2}$ Тогда $\operatorname{Var} X = 2\frac{q^2}{p^2} - \frac{q^2}{p^2} + \frac{q}{p} = \frac{q}{p} + (1 + \frac{q}{p}) = \frac{q}{p^2}$

- $S_n \sim NB(r, p), r \in \mathbb{N}$
  $X_1, \ldots, X_n$ – независимые, $X_i \sim Geom(p)$
  Тогда $EX = \frac{rq}{p}, \operatorname{Var} S_n = \frac{rq}{p^2}$

Равномерные:

- $U[0, 1]$
  $EX = \frac{1}{2}$
  $EX^2 = \frac{1}{3}$
  $\operatorname{Var} X = \frac{1}{12}$

- $U[a, b]$
  $EX = \frac{a+b}{2}$
  $\operatorname{Var} X = \frac{(b-a)^2}{12}$

- $N(0, 1)$
  $EX = 0$
  $EX^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} xxe^{-\frac{x^2}{2}} \,\mathrm{d}\, x = -\frac{1}{\sqrt{2\pi}} xe^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \,\mathrm{d}\, x = 1$

- $N(\mu, \sigma^2)$
  $X = \sigma N(0,1) + \mu$
  Тогда $EX = \mu, \operatorname{Var} X = \sigma^2$

- $\Gamma(\alpha, \lambda)$
  $p(x) = \dfrac{x^{\alpha-1}\lambda^\alpha e^{-\lambda x}}{\Gamma(\alpha)}\mathbb{1}(x \geq 0)$
  $EX = \dfrac{\alpha}{\lambda}\displaystyle\int_0^\infty \dfrac{x^\alpha \lambda^{\alpha+1} e^{-\lambda x}}{\Gamma(\alpha+1)} = \dfrac{\alpha}{\lambda}$
  $EX^2 = \dfrac{\alpha(\alpha+1)}{\lambda^2}\displaystyle\int_0^\infty \dfrac{x^{\alpha+1}\lambda^{\alpha+2}e^{-\lambda x}}{\Gamma(\alpha+2)} = \dfrac{\alpha(\alpha+1)}{\lambda^2}$
  $\operatorname{Var} X = \dfrac{\alpha(\alpha+1)}{\lambda^2} - \dfrac{\alpha^2}{\lambda^2} = \dfrac{\alpha}{\lambda^2}$

- $Exp(\lambda) = \Gamma(1, \lambda)$
  $EX = \dfrac{1}{\lambda}, \operatorname{Var} X = \dfrac{1}{\lambda^2}$

- $Cauchy(0,1)$
  $EX = \displaystyle\int_{-\infty}^{+\infty} \dfrac{1}{\pi}\dfrac{x}{1+x^2}\,\mathrm{d}\,x \neq 0$ – потому что интеграл расходится
  Матожидания не существует

Определение (мода)
Дискретный случай: $\operatorname{mod} X = x_k \Leftrightarrow P(X = x_k)$ – наибольшая
Непрерывный случай: $\operatorname{mod} X = x_* \Leftrightarrow x_* = \operatorname{argmax} p(x)$
Унимодальность – плотность имеет одну точку локального максимума
Бимодальность – плотность имеет две точки локального максимума (но мода все равно одна)
Определение (квантиль)
$\alpha \in [0, 1]$
q_α – квантиль порядка $\alpha \Leftrightarrow P(X \geq q_\alpha) \geq 1 - \alpha, P(X \leq q_\alpha) \geq \alpha$
Замечание
F строго монотонна $\Rightarrow F(q_\alpha) = \alpha, F^{-1}(\alpha) = q_\alpha$
F – не строго монотонна
$F^{-1}(\alpha) := \left[\begin{array}{l} \sup\{x : F(x) \leq \alpha\} \\ \inf\{x : F(x) \leq \alpha\} \end{array}\right.$
Определение
$\operatorname{med} X = q_{\frac{1}{2}}$
Утверждение

$\operatorname{Var} X = \min\limits_{a\in\mathbb{R}} \underbrace{E(X-a)^2}_{f(a)} \Leftrightarrow EX = \operatorname{argmin}_{a\in\mathbb{R}} E(X-a)^2$
$f(a) = EX^2 - 2aEX + a^2$
$a_{\min} = EX$
Утверждение
$\operatorname{med} X = \operatorname{argmin}_{a\in\mathbb{R}} E|X-a|$
Доказательство
Н.у.о. $\operatorname{med} X = 0$
Проверим, что $E|X-c| \geq E|X|$

$$c \geq 0 : |X-C| - |X| = \begin{cases} -c, & X > c \\ c-2X, & 0 < X \geq c \\ c, & X < 0 \end{cases}$$

$E(|X-c|-|X|) =$
$E((|X-c|-|X|)\mathbb{1}(X>c)) + E((|X-c|-|X|)\mathbb{1}(0<X\geq c)) + E((|X-c|-|X|)\mathbb{1}(X \geq 0)) =$
$-cE\mathbb{1}(X>c) + E((|X-c|-|X|)\mathbb{1}(0<X\geq c)) + cE(\mathbb{1}(X\leq 0)) \geq$
$-cP(X>c) - cP(0<X\leq c) + cP(X\leq 0) =$
$c(P(X\leq 0) - P(X>0)) = c(2P(X\geq 0)-1) \geq 0$

Определение
EX^k – k-ый момент
$E|X|^k$ – k-ый абсолютный момент
$E(X-EX)^k$ – k-ый абсолютный центральный момент
$E(X-EX)^k$ – k-ый центральный момент
$\gamma = \dfrac{E(X-EX)^3}{\sigma^3}$ – коэффициент асимметрии
$\gamma = 0$ – график симметричный
$\gamma > 0$ – правый «хвост» длиннее левого
$\gamma < 0$ – левый «хвост» длиннее правого
$\beta = \dfrac{E(X-EX)^4}{\sigma^4} - 3$ – коэффициент эксцесса
(константа 3 для нормировки: $\beta_{N(0,1)} = 0$)
$\beta > 0$ – плотности более сконцентрированы в центре
$\beta < 0$ – плотности более распределены

13 Числовые характеристики случайных величин

Определение
$\operatorname{Cov}(X, Y) = E((X - EX)(Y - EY))$
Свойства

1. $\operatorname{Cov}(X, Y) = E(XY) - EXEY$
2. Билинейность
3. $\operatorname{Cov}(X, c) = 0$
4. $\operatorname{Cov}(X, X) = \operatorname{Var} X$
5. X, Y – независимые $\Rightarrow \operatorname{Cov}(X, Y) = 0$
   Пример
   $X \sim N(0,1), Y = X^2$
   $\operatorname{Cov}(X, Y) = EX^2 - EXEX^2 = 0$, но X, Y не независимые

Определение
коэффициент корреляции $\rho(X, Y) = \dfrac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var} X \operatorname{Var} Y}}$
Замечание
X – нерожденная
$P(Y = c) = 1$
$P(X \in B, Y = 1) = P(X \in B)P(Y = 1)$
$P(X \in B, Y = 0) = 0 = P(X \in B)P(Y = 1)$
Тогда они независимые
Тогда пусть $\rho(X, Y) = 0$
Теорема

- $\rho(X, Y) \in [-1, 1]$
- $|\rho(X, Y)| = 1 \Leftrightarrow Y = aX + b, \operatorname{sign} a = \operatorname{sign} \rho(X, Y)$

Доказательство 1
$\rho(X, Y) = \dfrac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var} X \operatorname{Var} Y}} = \dfrac{\operatorname{Cov}(X - EX, Y - EY)}{\sqrt{\operatorname{Var} X \operatorname{Var} Y}} = \operatorname{Cov} \widetilde{X}, \widetilde{Y}$
$\widetilde{X} = \dfrac{X - EX}{\sqrt{\operatorname{Var} X}}$

$\widetilde{Y} = \dfrac{Y - EY}{\sqrt{\operatorname{Var} Y}}$
$E\widetilde{X} = 0 = E\widetilde{Y}$
$\operatorname{Var}\widetilde{X} = \dfrac{\operatorname{Var}\widetilde{Y}}{(\sqrt{\operatorname{Var}\widetilde{Y}})^2} = 1$
$\operatorname{Var}(\widetilde{X} + \widetilde{Y}) = \operatorname{Var}\widetilde{X} + \operatorname{Var}\widetilde{Y} + 2\operatorname{Cov}(\widetilde{X}, \widetilde{Y}) = 2(1 + \rho(X, Y)) \geq 0$
$\operatorname{Var}(\widetilde{X} - \widetilde{Y}) = 2(1 - \rho(X, Y)) \geq 0$
Доказательство 2($\Rightarrow$)
$\rho(X, Y) = 1 \Rightarrow \operatorname{Var}(\widetilde{X} - \widetilde{Y}) = 0 \Rightarrow \widetilde{X} - \widetilde{Y} = c$
$\rho(X, Y) = -1 \Rightarrow \widetilde{X} + \widetilde{Y} = c$
Доказательство 2($\Leftarrow$)
$a \neq 0 : \rho(aX + b, X) = \dfrac{\operatorname{Cov}(aX + b, X)}{\sqrt{\operatorname{Var}(aX + b)\operatorname{Var} X}} = \dfrac{a\operatorname{Cov} X, X}{\sqrt{a^2 \operatorname{Var}^2 X}} = \dfrac{a}{|a|} = \operatorname{sign} a$
$a = 0$ – очевидно
Определение
$X = (X_1, \ldots, X_n)^T$ – случайный вектор
$EX = (EX_1, \ldots, EX_n)^T$ – вектор математических ожиданий
Замечание
X – случайная матрица
$EX = (EX_{ij})$ – матрица математических ожиданий
Свойства
$E(AX + b) = AEX + b$ – линейность
Определение
$X = (X_1, \ldots, X_n)^T$ – случайный вектор
$\operatorname{Var} X = (\operatorname{Cov}(X_i, X_j))_{i,j} = E((X - EX)(X - EX)^T)$
Свойства

1. $\operatorname{Var} X = (\operatorname{Var} X)^T$

2. $\operatorname{Var} X \geq 0$, т.е. $\forall\, t \in \mathbb{R}^n t^T \operatorname{Var} X t \geq 0$
   Доказательство
   $t^T \operatorname{Var} X t = t^T(E(X - EX)(X - EX)^T)t = Et^T(X - EX)(X - EX)^T t = E(t^T(X - EX))^2 \geq 0$

3. $\operatorname{Var}(AX + b) = A\operatorname{Var}(X)A^T$

4. X, Y – независимые $\Rightarrow \operatorname{Var} X \pm Y = \operatorname{Var} X + \operatorname{Var} Y$
   Доказательство

$$\mathrm{Cov}(X_i \pm Y_i, X_j \pm Y_j) = \mathrm{Cov}(X_i, X_j) \underbrace{\pm \mathrm{Cov}(X_i, Y_j) \pm \mathrm{Cov}(X_j, Y_i)}_{0} + \mathrm{Cov}(X_j, Y_j)$$

Пример (гауссовский вектор)
$X \sim N(\mu, \Sigma), \mu \in \mathbb{R}^n, \Sigma = \Sigma^T > 0$
Тогда $X = \sqrt{\Sigma}Y + \mu, Y \sim N(0, E)$ – стандартный гауссовский вектор
$Y = (Y_1, \ldots, Y_n)$ – стандартный гауссовский вектор $\Leftrightarrow Y_i \sim N(0,1), Y_i$ – независимые
$EY = 0$
$\operatorname{Var} Y = E_n$
$EX = \sqrt{\Sigma}EY + \mu = \mu$ – вектор математических ожиданий
$\operatorname{Var} X = \sqrt{\Sigma} \operatorname{Var} Y (\sqrt{\Sigma})^T = \Sigma$
Пример
X – гауссовский вектор
Покажем, что X_i, X_j – независимые $\Leftrightarrow \mathrm{Cov}(X_i, X_j) = 0 \forall\, i, j$
$i := 1, j := 2$

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sqrt{\Sigma_{12}} & 0 \\ 0 & \ddots \end{pmatrix} \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \sqrt{\Sigma_{12}} & \ldots \end{pmatrix} \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \Rightarrow (X_1, X_2)^T \sim N((\mu_1, \mu_2)^T, \Sigma_{12})$$

$(X_1, X_2)^T \sim N((\mu_1, \mu_2)^T, \mathrm{diag}(\sigma_1^2, \sigma_2^2))$
$\sigma_1^2 = \operatorname{Var} X_1$
$\sigma_2^2 = \operatorname{Var} X_2$

$$p(X_1, X_2) = \frac{\exp(-\frac{1}{2}(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}))}{\sqrt{2\pi}^2 \sigma_1 \sigma_2}$$

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \sqrt{\Sigma}_X & \sqrt{\Sigma}_{XY} \\ \sqrt{\Sigma}_{XY} & \sqrt{\Sigma}_Y \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} + \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}$$

$X = \sqrt{\Sigma}_X U_1 + \sqrt{\Sigma}_{XY} U_2 + \mu_X$

$$\operatorname{Var} X = \sqrt{\Sigma}_X \quad \sqrt{\Sigma}_{XY} \begin{matrix} \sqrt{\Sigma}_X \\ \sqrt{\Sigma}_{XY} \end{matrix} = \sqrt{\Sigma_X}^2 + \sqrt{\Sigma_{XY}}^2$$

$$\begin{pmatrix} \sqrt{\Sigma}_X & \sqrt{\Sigma}_{XY} \\ \sqrt{\Sigma}_{XY} & \sqrt{\Sigma}_Y \end{pmatrix} \begin{pmatrix} \sqrt{\Sigma}_X & \sqrt{\Sigma}_{XY} \\ \sqrt{\Sigma}_{XY} & \sqrt{\Sigma}_Y \end{pmatrix} = \begin{pmatrix} \sqrt{\Sigma_X}^2 + \sqrt{\Sigma_{XY}}^2 & \sqrt{\Sigma}_{XY}(\sqrt{\Sigma}_X + \sqrt{\Sigma}_Y) \\ \sqrt{\Sigma}_{XY}(\sqrt{\Sigma}_X + \sqrt{\Sigma}_Y) & \sqrt{\Sigma_Y}^2 + \sqrt{\Sigma_{XY}}^2 \end{pmatrix}$$

14 Условное распределение числовых характеристик

Определение

(X, Y) – случайный вектор

Дискретный случай: $P(X = x_i | Y = y_j) = \dfrac{P(X = x_i, Y = y_j)}{P(Y = y_j)}$

$P(Y = y_j) = \sum\limits_k P(Y = y_j | X = x_i) P(X = x_i)$

$P(X = x_i | Y = y_j) = \dfrac{P(X = x_i, Y = y_j)}{\sum_k P(Y = y_j | X = x_i) P(X = x_i)}$ – теорема Байеса

Непрерывный случай: $p_{X|Y}(x|y) = \dfrac{p_{X,Y}(x, y)}{p_Y(y)}$

$p_Y(y) = \int\limits_{\mathbb{R}} p_{Y|X}(y|x) p_X(x) \,\mathrm{d}\, x$

$p_{X|Y}(x|y) = \dfrac{p_{X,Y}(x, y)}{\int_{\mathbb{R}} p_{Y|X}(y|x) p_X(x) \,\mathrm{d}\, x}$ – теорема Байеса

Определение

$$E(X|Y = y) = E(X|Y) = \begin{cases} \sum\limits_i x_i P(X = x_i | Y = y) \\ \int\limits_{\mathbb{R}} x p(x|y) \,\mathrm{d}\, x \end{cases}$$

Условное математическое ожидание – случайная величина

Свойства

1. $E(aX + bY|U) = aE(X|U) + bE(Y|U)$
2. $E(X|Y) = \operatorname{argmin}_{f(Y)} E(X - f(Y))^2$
3. $E(X|X) = X$
4. X, Y – независимые $\Rightarrow E(X|Y) = E(X)$
5. $EX = E_Y(E_{X|Y}(X|Y))$

   Доказательство

   $EX = \sum\limits_k x_k P(X = x_k) = \sum\limits_k x_k \sum\limits_j P(X = x_k | Y = y_j) P(Y_j) = \sum\limits_j P(Y = y_j) \sum\limits_k x_k P(X_k | Y = y_k) = \sum\limits_j P(Y = y_j) E(X | Y = y_j) = E_Y(E(X|Y))$

Пример
$(X,Y) \sim N(\begin{pmatrix}\mu_X \\ \mu_Y\end{pmatrix}, \begin{pmatrix}\sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2\end{pmatrix})$
Тогда $E(Y|X) = aX + b$
$E(Y|X) = E(Y - Ax + AX|X) = E(Y - aX|X) + E(aX|X) = E(Y - \frac{\sigma_{XY}}{\sigma_X^2}) + \frac{\sigma_{XY}}{\sigma_X^2}X = (X - \mu_X)\frac{\sigma_{XY}}{\sigma_X^2} + \mu_X$
$\begin{pmatrix}Y - aX \\ X\end{pmatrix} \sim N(\cdot,\cdot)$
$0 = \operatorname{Cov}(Y - aX, X) = \operatorname{Cov}(Y, X) - a\operatorname{Cov}(X,X) \Rightarrow a = \frac{\sigma_{XY}}{\sigma_X^2}$
Определение
$\operatorname{Var}(X|Y) = E((X - E(X|Y))|Y) = E(X^2|Y) - (E(X|Y))^2$
$\operatorname{Var} X = \operatorname{Var} E(X|Y) + E\operatorname{Var}(X|Y)$

15 Сходимости

$X_n \xrightarrow{a.s.} X \Leftrightarrow P(\{\omega : X_n(\omega) \to X(\pi)\}) = 1$ – почти наверное
$X_n \xrightarrow{p} X \Leftrightarrow \forall\varepsilon > 0\ P(|X_n - X| > \varepsilon) \to 0$ – по вероятности
$X_n \xrightarrow{Lp} \Leftrightarrow E|X_n - X|^p \to 0, \geq 1$
$X_n \xrightarrow{d} X \Leftrightarrow F_n(x) \to F(x), x \in C(F)$ – по распределению
$F_n \Rightarrow F \Leftrightarrow \forall f$ – непр. $\displaystyle\int f\,\mathrm{d}\,F_n(x) \to \int f\,\mathrm{d}\,F(x)$ – слабая
Теорема
$X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{d} X$
$X_n \xrightarrow{Lp} X \Rightarrow X_n \xrightarrow{d} X$

$\mathcal{F}$ – множество функций распределения
$\mathcal{G} = \{G : \mathbb{R} \to \mathbb{R}|G$ – монот. возр, непр. справа, $G(-\infty) \geq 0, G(+\infty \leq 1)\}$
– расширенные распределения
Распределения:
$P(X \in \mathbb{R}) = G(+\infty) - G(-\infty)$
$P(X = -\infty) = G(-\infty)$
$P(X = +\infty) = 1 - G(+\infty)$
$P(X_n = n) = P(X_n = -n) = \frac{1}{2}$

$$F_n(x) = \begin{cases} 0, & x < -n \\ \frac{1}{2}, & -n \le x < n \\ 1, & x \ge n \end{cases}$$
$F_n(x) \xrightarrow{d} \frac{1}{2}$

$F_n(x) \not\Rightarrow \frac{1}{2}$

Замечание

1. $\xrightarrow{d} \Rightarrow$ можно задать на $\mathcal{G}$

2. $G_n, G \in \mathcal{G}$ $(G_n \Rightarrow G) \Rightarrow G_n \xrightarrow{d} G$
   В обратную сторону не действует

Теорема (Хелли)
$G_n \in G \Rightarrow \exists\, G_{n_k}, G \in G : G_{n_k} \Rightarrow G$
(без доказательства)
Следствие
Если всякая слабо сходящаяся подпоследовательность сходится к одной и той же точке G, то $G_n \Rightarrow G$
Определение (плотные распределения)
$F_n \in \mathcal{G}$
$\{F_n\}$ – плотная, если $\forall\, \varepsilon > 0 \exists\, M : \inf\limits_n \underbrace{(F_n(M) - F_n(-M-0))}_{P(M \le X_n \le M)} > 1 - \varepsilon$

Т.е. «хвосты» равномерно маленькие
Определение
Пусть L – подмножество непрерввных и ограниченных функций
L – определеяет распределение $\Leftrightarrow$ $F \in \mathcal{F}, G \in \mathcal{G}, \forall\, f \in L$ $\displaystyle\int f \,\mathrm{d}\, F = \int f \,\mathrm{d}\, G \Rightarrow F = G$
Теорема (Критерий существования слабого предела из $\mathcal{F}$)
$F_n \in \mathcal{F}, L$ – определяет распределение Тогда $\exists\, F \in \mathcal{F} : F_n \Rightarrow F \Leftrightarrow$

1. $\{F_n\}$ – плотное

2. $\exists \lim\limits_n \displaystyle\int f \,\mathrm{d}\, F_n \forall\, f \in L$

Доказательство $\Rightarrow$
2 выполнено
$$\int f \,\mathrm{d}\, F_n \to \int f \,\mathrm{d}\, F \;\forall f$$
$F_n \xrightarrow{d} F$
$\exists\, M > 0 : M, -M \in C(F), F_n(+M) - F_n(-M-0) \to F(M) - F(-M) > 1 - \varepsilon$
Доказательство $\Leftarrow$
$\exists\, F_{n_k} : F_{n_k} \Rightarrow G \in \mathcal{G}$
Тогда $F_{n_k} \xrightarrow{d} G$
$F_{n_j} \Rightarrow F \in \mathcal{F}$
$$\exists \lim \int f \,\mathrm{d}\, F_n$$
$$\int f \,\mathrm{d}\, F_{n_k} \Rightarrow A$$
$$\int f \,\mathrm{d}\, F_{n_k} \Rightarrow \int f \,\mathrm{d}\, G$$
$$\int f \,\mathrm{d}\, F_{n_j} \Rightarrow A$$
$$\int f \,\mathrm{d}\, F_{n_j} \Rightarrow \int f \,\mathrm{d}\, F$$
Тогда $G = F$
Следствие
L – определяет распределение
$$\int f \,\mathrm{d}\, F_n \to f \,\mathrm{d}\, F$$
Пусть верно одно из трех условий

1. $\{F_n\}$ – плотное
2. $F \in \mathcal{F}$
3. $1 \in L$

Тогда $F_n \Rightarrow F \in \mathcal{F}$
Доказательство 1
По теореме
Доказательство 2
Тогда выполнено 1
Доказательство 3

$\underbrace{\int \mathrm{d}\, F_n}_{1} \to \int \mathrm{d}\, F$

Тогда $\int \mathrm{d}\, F = 1$

$\int \mathrm{d}\, F = F(+\infty) - F(-\infty)$

Тогда $F(+\infty) = 1, F(-\infty) = 0$

Тогда выполнено условие 2

Теорема

Пусть $F_n \xrightarrow{d} F \in \mathcal{F}$

Т.е. $F_n(x) \to F(x)\ \forall\, x \in C(F)$

Покажем, что $\forall\, f \in \widetilde{C}_1^0(\mathbb{R})$ $\int_{-\infty}^{\infty} f \,\mathrm{d}\, F_n(x) \to \int_{-\infty}^{\infty} f \,\mathrm{d}\, F(x)$

Доказательство

$\int_{-\infty}^{\infty} f \,\mathrm{d}\, F(x) = \underbrace{fF\Big|_{-\infty}^{\infty}}_{0} - \int_{-\infty}^{\infty} F \,\mathrm{d}\, f(x)$

$\int_{-\infty}^{\infty} f \,\mathrm{d}\, F_n(x) = \underbrace{fF_n\Big|_{-\infty}^{\infty}}_{0} - \int_{-\infty}^{\infty} F_n \,\mathrm{d}\, f(x)$

Тогда $F_n \Rightarrow F$