Теория Вероятностей

Aleksandr Shirokov
improfeo@yandex.ru

Saint-Petersburg
2020

Содержание

1 Введение в теорию вероятностей

1.1 События

Опр: **Пространством элементарных событий** Ω называется множество, содержащее все возможные результаты случайного эксперимента, из которых в эксперименте происходит ровно один. Элементы ω этого множества называют элементарными исходами.

Опр: **Событиями** называются подмножества множества Ω. Говорят, что в результате эксперимента произошло событие $A \subseteq \Omega$, если в эксперименте произошел один из элементарных исходов, входящих в множество A.

Опр: **Достоверным** называется событие, которое обязательно происходит в результате эксперимента, т.е. единственное событие, включающее все элементарные исходы — событие Ω.

Опр: **Невозможным** называется событие, которое не может произойти в результате эксперимента, т.е. событие, не содержащее ни одного элементарного исхода. Всегда $\emptyset \subset \Omega$.

Опр: **Суммой** $A + B$ или объединением $A \cup B$ событий AA и BB называется событие, состоящее в том, что произошло либо событие A, либо событие B, либо оба события одновременно, т.е. множество $A + B$ содержит элементарные исходы, входящие либо в множество A, либо в B или в оба.

Опр: **Умножением** AB или пересечением $A \cap B$ событий A и B называется событие, состоящее в том, что произошли оба события A и B одновременно, т.е. множество AB содержит элементарные исходы, входящие одновременно и в множество A, и в множество B.

Опр: **Дополнением** $A \backslash B$ события B до A называется событие, состоящее в том, что произошло событие A, но не произошло B, т.е. множество $A \backslash B$ содержит элементарные исходы, входящие в множество A, но не входящие в B.

Опр: **Противоположным** к событию А называется событие $\bar{A} = \Omega \backslash A$, состоящее в том, что событие A в результате эксперимента не произошло, т.е. множество $\bar{A}$ состоит из элементарных исходов, не входящих в А.

Опр: Говорят, что несколько событий образует **полную группу событий**, если в результате опыта непременно должно появиться хотя бы одно из них.

Опр: События A и B называются **несовместными**, если их совместное появление невозможно, т.е. $AB = \emptyset$

Опр: Будем говорить, что событие A **содержится** в B, если все элементарные события, содержащиеся в A, содержатся и в B: $A \subset B$

1.2 Вероятности на дискретных пространствах

Опр: будем говорить, что Ω заданы **верояности элеметарных событий**, если задана функция $P : \Omega \to \mathbb{R}$, такая, что

$$\forall \omega_i \quad P(\omega_i) \geq 0 \quad \sum_{\Omega} p(\omega_i) = 1$$

Опр: **вероятностью события** A называется сумма вероятностей элементарных событий, входящих в A:

$$p(A) = \sum_{i:\omega_i \in A} p(\omega_i)$$

Свойства вероятностей событий:

1. $p(\emptyset) = 0$
2. $p(\Omega) = 1$
3. $p(\bar{A}) = 1 - p(A)$
4. если $A \subset B$, то есть наступление события A влечет наступление события B, то $p(A) \leq p(B)$
5. $0 \leq p(A) \leq 1$
6. если $A \subset B$, заключающегося в наступлении события B при одновременнои наступлении события A равна $P(B \backslash A) = p(B) - p(A)$
7. Теорема сложения вероятностей: вероятность наступления хотя бы одного из произвльных событий (не обязательно невосместных) равна: $P(A \cup B) = p(A) + p(B) - p(A \cap B)$
8. $p(A \cap B) = 0 \Rightarrow P(A \cup B) = p(A) + p(B)$

1.3 Схема равновозможных исходов

1.4 Геометрическая вероятность

1.5 Алгебры и сигма-алгебры

1.6 Аксиоматическое построение вероятности

1.7 Условные вероятности

Опр: условной вероятностью A при условии того, что событие B произошло (считается, что $P(B) > 0$), называется

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

Формула умножения вероятности:

$$p(A \cap B) = p(B) \cdot p(A|B) = p(A) \cdot p(B|A)$$

$$p(A_1 \cap A_2 \cap \ldots A_n) = p(A_1) \cdot p(A_2|A_1) \cdot p(A_3|A_1 \cap A_2) \ldots \cdot p(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

Свойства условной вероятности

1.8 Формула полной вероятности

Формула полной вероятности:

Вероятность события A, которое может наступить лишь при появлении одного из несовместных событий $H_1, H_2, \ldots, H_n$, образующих полную группу событий, равна

$$p(A) = \sum_{i=1}^{n} p(H_i) \cdot p(A|H_i)$$

,

где $\sum_{i=1}^{n} p(H_i) = 1$

1.9 Теорема Байеса

Теорема Байеса:

Пусть $\{H_i\}$ - разбиение Σ. Тогда:

$$p(H_i|A) = \frac{p(A \cap H_i)}{p(A)} = \frac{p(A \cap H_i)}{\sum_{i=1}^{n} p(H_i) \cdot p(A|H_i)}$$

1.10 Независимость

Опр: Событие A называется независимым от события B, если вероятность события A не зависит от того, произошло событие B или нет. Событие A называется зависимым от события B, если вероятность события A меняется в зависимости от того, произошло событие B или нет.

$$p(A|B) = p(A)$$

Замечание: если A не зависит от B, то и B не зависит от A.

$$p(B|A) = p(B)$$

Опр: будем говорить, что A и B независимы, если

$$p(A \cap B) = p(A) \cdot p(B)$$

Замечание: A и B независимые $\Rightarrow$ $\bar{A}$ и B независимы

1.11 Схема Бернулли

Опр: схемой Бернулли назовём последовательность независимых испытаний, в каждом из которых может наступить лишь один исход.

Теорема: пусть в схеме Бернулли каждое событие наступает с вероятностью p, и не наступает с вероятностью $1-p$. Обозначим за $P_n(k)$ - событие A наступило k раз в n испытаниях. Тогда вероятность данного события равна:

$$P_n(k) = C_n^k p^k q^{n-k} = \frac{n!}{k! \cdot (n-k)!} \cdot p^k q^{n-k}$$

Теорема: Пусть k_1 и k_2 - целые числа, $0 \le k_1 \le k_2 \le n$. Обозначим за $P_n(k_1, k_2)$ - вероятность того, что событие A настуипло не менее m_1 и не более m_2 раз в n испытаниях (количество успехов от k_1 до k_2 в серии из n испытаний. Тогда:

$$P_n(k_1, k_2) = \sum_{j=k_1}^{k_2} C_n^k p^j q^{n-j}$$

Теорема Пуассона:

Предположим, что произведение $\lambda = np$ является постоянной величиной, когда n неограниченное возрастает. Тогда вероятность возникновения k успехов в серии из n испытаний равна:

$$\lim_{n\to\infty} p_n(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Локальная Теорема Муавра-Лапласа:

$$P_n(k) = C_n^k p^k q^{n-k} \to \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \cdot e^{-\frac{(k-np)^2}{2npq}}, n, k \to \infty$$

Интегральная Теорема Муавра-Лапласа:

Пусть $P_n(k_1, k_2)$ - вероятность, что в серии из n испытаний количество успехов от k_1 до k_2. Обозначим за $a_1 = \frac{k_1 - np}{\sqrt{npq}}, a_2 = \frac{k_2 - np}{\sqrt{npq}}$

Тогда:

$$\lim_{n\to\infty} \left| P_n(k_1, k_2) - \frac{1}{\sqrt{2\pi}} \int_{a_1}^{a_2} e^{-\frac{t^2}{2}} dt \right| = 0$$

Положим

$$N(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

Данная функция называется функцией Лапласа.

2 Случайные величины

2.1 Случайные величины и способы их описания

Рассмотрим сначала случайный эксперимент с дискретным пространством исходов, т. е. с таким пространством влементарных событий $\Omega = \{\omega_1, \omega_2, \ldots\}$, которое состоит из конечного или счетного числа исходов. Пусть есть величина, которая в результате случайного эксперимента принимает различные числовые значения в зависимости от наступления того или иного исхода, при этом каждому исходу соответствует только одно число. Иными словами, на пространстве элементарных событий задана функция.

Опр: функция, заданная на пространстве элементарных событий $\Omega = \{\omega_1, \omega_2, \ldots\}$, называется случайной величиной ξ. Для любого исхода ω значение $x = X(\omega)$ - это реализация случайной величины при данном исходе.

В ходе случайного эксперимента реализуется лишь один какой-то исход, это означает, что в результате эксперимента наблюдается лишь какое-то одно значение случайной величины (одна реализация) из всех возможных.

Опр: случайой величиной называется функция $\xi = \xi(\omega)$, заданная на пространстве элементарных событий Ω, для которой событие $\{\xi < x\} = \{\omega : \xi(\omega) < x\}$ принадлежт σ-алгебе A для любого вещественного x.

$\{\xi < x\}$ - дает рассматривать вероятности событий, поскольку вероятности определеные только на множествах из A. Кроме того через события вероятости и $x \in (-\infty, \infty)$ с помощью известных операция над событиями можно выразить сколь угодно сложное событие, связанное со случайной величиной ξ

Вся совокупность вероятностей $P(\xi < x)$, $x \in (-\infty, \infty)$ задает закон распредлеения случайной величины ξ. Часто для краткости закон распределения называют распределением случайной величины.

Опр: распределением случайной величины ξ назовём $P(\xi \in B)$ для $B \in \mathbb{B}$ такую, что:

$$P(\xi \in [a; b]) = P(\xi < b) - P(\xi < a)$$

Опр: функцией распределения случайной величины ξ назовём:

$$F_\xi(x) = P(\omega : \xi(\omega) < x) = P(\xi(\omega) < x) = P(\xi < x)$$

Свойства функции распределения:

- $F_\xi(x)$ - монотонно неубывающая: если $x_1 < x_2$, тогда $F(x_1) \leq F(x_2)$
- $F_\xi(x)$ непрерывна слева (непрерывности справа нет)
- $\lim_{x\to\infty} F_\xi(x) = F_\xi(\infty) = 1$
- $\lim_{x\to-\infty} F_\xi(x) = F_\xi(-\infty) = 0$

- $P(\omega : a \le \xi(\omega) \le b) = F(b) - F(a)$

ξ называется дискретной случайной величиной, если она принимает счетное число значений. Если множество значений конечно, то случайная величина называется простой.

Законом распределения дискретной случайной величины называется совокупность всех возможных значений x_i и соответствующих им вероятностей $p_i = P(\omega : \xi = x_i)$, таких, что

$$\sum_{i=1}^{\infty} P(\xi = x_i) = 1$$

Опр: будем говорить, что случайная величина имеет абсолютное непрерывное распределение, если найдетя функция $f_\xi(x) \ge 0$, что $\forall x \in \mathbb{R}(-\infty, \infty)$:

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

$f_\xi(x)$ называется плотностью распределения случайной величины.

Если случайная величина ξ имеет плотность распределения, то ее функция распределения непрерывна, поскольку интеграл - непрерывная функция верхнего предела.

$$P(\omega : a \le \xi(\omega) \le b) = F(b) - F(a) = \int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{a} f(t)dt = \int_{a}^{b} f(t)dt$$

Свойства плотности вероятности:

1. $\int\limits_{-\infty}^{\infty} f(t)dt = 1$ - площад под кривой плотности распределения равна единице
2. $f(x) = F'(x)$, если функция распределения дифференцируема
3. $P(\xi = x) = \int_x^x f(t)dt = 0$

Пусть $[x; x+\delta x)$ - интервал произвольно малой длины. Вероятность попадания случайной величины в этот интервал приближенно равна проивезеднию значения плотности распределения в точке x на длину этого интервала. То есть пропорциональна длине интервала.

Выведем плотность распределения. $F_\xi(t)$ и $f_\xi(t)$ связаны с друг другом соотношением:

$$F_\xi(t) = \int\limits_0^t f_\xi(x)dx$$

$$f_\xi(t) = \frac{\partial F_{\xi(t)}}{dt} = \lambda \cdot e^{-\lambda t}$$

2.2 Совместные функции распределения нескольких случайных величин

Для независимости случайных величин необходимо и достаточно, чтобы $\forall(x_1, x_2, \dots, x_n)$:

$$F_{\xi_1,\xi_2,\dots,\xi_n}(x_1, x_2, \dots, x_n) = F_{\xi_1}(x_1) \cdot F_{\xi_2}(x_2) \dots F_{\xi_n}(x_n)$$

Кроме того, если $\xi_1, \xi_2, \dots, \xi_n$ - дискретно, то необходимое и достаточное условие независисмости:

$$P(\xi_1 = x_1, \dots, \xi_n = x_n) = P(\xi_1 = x_1) \cdot \dots \cdot P(\xi_n = x_n)$$

Опр: если существует такая неотрицательная функция $f(t_1, t_2, \dots, t_k)$, что функция распределения $F(x_1, x_2, \dots, x_k)$ для всех $x_1, x_2, \dots, x_k$ представима в виде

$$F(x_1, x_2, \dots, x_k) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_k} f(t_1, t_2, \dots, t_k) dt_1 dt_2 \dots dt_k$$

то $f(t_1, t_2, \dots, t_k)$ называется совместной плотностью распредлеения случайных величин $\xi_1, \xi_2, \dots, \xi_k$

Свойства совместной плотности распределения:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(t_1, t_2, \dots, t_k) dt_1 dt_2 \dots dt_k = 1$$

и совместная плотность распределения есть частные производные от совместной функции распределения.

Если существует совместная плотность, значит вектор непрерывен и непрерывность выражается как равенство совместной плотности распределения и произведения отдельных ее плотностей.

2.3 Характеристики случайных величин

2.3.1 Математическое ожидание (Expected Value)

Математиическое ожидание — одно из важнейших понятий в теории вероятностей, означающее среднее (взвешенное по вероятностям возможных значений) значение случайной величины. В случае непрерывной случайной величины подразумевается взвешивание по плотности распределения (более строгие определения см. ниже). Математическое ожидание случайного вектора равно вектору, компоненты которого равны математическим ожиданиям компонент случайного вектора.

Опр: математическим ожиданием дискретной случайной величины, принимающей значения $x_1, \ldots, x_n$ с вероятностями $p_1, \ldots, p_n$, $P(\xi = x_i) = p_i$ назовём

$$E\xi = \sum_{i=1} x_i p_i$$

если данный ряд абсолютно сходится.

Опр: математическим ожиданием случайной величины ξ с непрерывным распределением с плотностью распределения $f_{xi}(x)$ называется:

$$E\xi = \int_{-\infty}^{\infty} x \cdot f_{\xi}(x) dx$$

если данный интеграл абсолютно сходится, в противном случае говорят, что математическое ожидание не существует.

В общем случае математическое ожидание представляет собой несобственный интеграл Стилтьеса по функции распределения $E\xi = \int x dF_{\xi}(x)$, если он существует.

Свойства математического ожидания:

1. Для любой борелевской функции $h : \mathbb{R} \to \mathbb{R}$

   - $Eh(\xi) = \sum_{i=1}^{\infty} h(x_i) p_i$
   - $Eh(\xi) = \int_{-\infty}^{\infty} h(x) f(x) dx$, если распределение непрерывно

2. Математическое ожидание ξ существует тогда и только тогда, когда $E|\xi| < \infty$

3. $E\xi = c$, если $\xi = c$ (рассматривая $h(x) = c$)

4. $E(c\xi) = c \cdot E\xi$

5. $E(a\xi + b\eta) = aE\xi + bE\eta$

6. Математическое ожидание суммы любых случайных величин равно сумме их математических ожиданий: $E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2$

7. если ξ и η - независимы, то $Eg(\xi)h(\eta) = Eg(\xi) \cdot Eh(\eta)$

8. если $\xi \geq 0$ почти навероноее ($P\{\xi \geq 0\} = 1$), тогда $E\xi \geq 0$

9. если $\xi \geq \eta$ почти наверное, то $E\xi \geq E\eta$

10. $|E\xi| \leq E(|\xi|)$

11. $E(\sum_{i=1}^{k} \xi_i) = \sum_{i=1}^{k} E(\xi_i)$

2.3.2 Дисперсия и отклонения

Опр: средним отклонением случайной величины назовём $E|\xi - E\xi|$.

Опр: дисперсией случайной величины назовём $D\xi = E(\xi - E\xi)^2$, если она существует. Смысл дисперсии хаключается в том, что она характеризует средний квадратичный разброс случайной величины вокруг своего математического ожидания.

Опр: стандартным отклонением (среднеквадратическим отклонением) случайной величины назовём $\sigma(\xi) = \sqrt{D\xi}$

Дисперсия $D\xi$ случайной величины ξ с дискретным распределением, заданным таблицей $P(\xi = x_i) = p_i$ выражается формулой:

$$D\xi = \sum_{i=1}^{\infty} (x_i - E\xi)^2 \cdot p_i$$

Дисперсия $D\xi$ случайной величины ξ с непрерывным распределением с плотностью распределения $f_\xi(x)$ вычисляется:

$$D\xi = \int_{-\infty}^{\infty} (x - E\xi)^2 \cdot f_\xi(x) dx$$

Свойства дисперсии:

1. $D(\xi) = 0, \xi = a(const)$

2. $D(\xi + a) = D\xi$

3. $D(a\xi) = a^2 \cdot D\xi$

4. если дисперсия существует, то $D\xi \geq 0$
5. $\sigma(a\xi) = |a|\sigma\xi$
6. если ξ и η независимы, то $D(\xi + \eta) = D\xi + D\eta$
7. если ξ и η независимы, то $D(\xi - \eta) = D\xi - D\eta$
8. $D\xi = E\xi^2 - (E\xi)^2$
9. $D(\sum_{i=1}^k \xi_i) = \sum_{i=1}^k D(\xi_i)$ - для независимых (!)
10. если дисперсия равна нулю, то в силу определения $P(\xi = E\xi) = 1$, т.е фактически ξ - детерминированная (не случайная) величина

Пусть $E|\xi|^k < \infty$

Опр: k-ым начальным моментом случайной величины ξ назовём $E\xi^k$

Опр: Число $E|\xi|^k$ называется абсолютным моментом порядка k случайной величины ξ.

Опр: Число $E(\xi - E\xi)^k$ называется центральным моментом порядка k случайной величины ξ.

Опр: Число $E|\xi - E\xi|^k$ называется абсолютным центральным моментом порядка k случайной величины ξ.

Если существует момент порядка $t > 0$ случайной величины ξ, то существуют и её моменты порядка s, где $0 < s < t$

Неравенство Коши-Буняковского:

Для любых случайных величин ξ и η выполняется неравенство:

$$E(|\xi\eta|) \leq (E\xi^2)^{\frac{1}{2}} \cdot (E\eta^2)^{\frac{1}{2}}$$

2.4 Ковариация, корреляция, квантиль

Опр: путь ξ, η - две случайные величины, определенные на одном и том же вероятностом пространстве. Тогда ковариацией случайых величин называется выражение вида:

$$cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$$

Свойства ковариации:

1. В силу линейности математического ожидания, ковариация может быть записана как:

$$cov(\xi, \eta) = E(\xi\eta) - E\xi E\eta$$

2. ковариация симметрична: $cov(\xi, \eta) = cov(\eta, \xi)$
3. $cov(a\xi + b, c\eta + d) = a \cdot c \cdot cov(\xi, \eta)$
4. $cov(\xi, a) = 0$
5. $cov(\xi, \eta) = 0$, если ξ, η - независимы (обратное неверно)
6. $cov(\xi, \xi) = D\xi$
7. аддитивность: $cov(\xi + \eta, \nu) = cov(\xi, \eta) + cov(\xi, \nu)$
8. переписанное неравенство Коши-Буняковского: $cov^2(\xi, \eta) \leq D\xi \cdot D\eta$
9. $D(\xi_1, \xi_2, \dots, \xi_n) = \sum_{i=1}^{n} D\xi_i + \sum_{i \neq j} cov(\xi_i, \xi_j) = \sum_{i,j} cov(\xi_i, \xi_j)$

Величина ковариации не является безразмерной. Иначе говоря при умножении случайных величин на какое-то число, ковариация тоже умножеается на это число. Но умножение на число не сказывается на степени зависимости величин (они не становятся от этого более зависимыми), так что большое значение ковариации не означает сильной зависимости. Это плохо.

Нужно как-то нормировать ковариацию, получив безразмерную величину, абсолютное значение которой не менялось бы при умножении случайных чисел на число и свидетельствовало о силе зависимости случайных величин.

Самая сильная зависимость - функциональная, а из функциональных - линейная.

Опр: коэффициентом корреляции случайных величин назовём

$$\rho(\xi, \eta) = \frac{cov(\xi, \eta)}{\sqrt{D\xi \cdot D\eta}}$$

Коэффициент корреляции характеризует степень линейной зависимости случайных величин. Чем больше по модулю коэффициент корреляции, тем сильнее линейная зависимость между величинами.

Опр: если ковариация или коэффициент корреляции равны нулю, то такие величины называются некоррелированными.

Свойства коэффициента корреляции:

1. для любых постоянных a, b, c, d:

   $$\rho(a\xi + b, c\eta + d) = sign(ac)\rho(\xi, \eta)$$

   - линейные преобразования случайных величин не изменяют степени их линейное независимости, при этом меняется лишь знак зависимости.

2. $-1 \leq \rho \leq 1$

3. $|\rho(\xi, \eta)| = 1$ тогда и только тогда, когда существуют такие постоянные a и b, что $\eta = a\xi + b$ с вероятностью единица

4. $D(\xi + \eta) = D\xi + D\eta + 2 \cdot cov(\xi, \eta)$

5. $\rho(\xi, \eta) = 0$, если величины ξ и η независимы

6. Говорят, что ξ и η отрицательно коррелированы, если $\rho(\xi, \eta) < 0$, положительно коррелированы, если $\rho(\xi, \eta) > 0$, и некоррелированы, если $\rho(\xi, \eta) = 0$

Квантили:

Опр: для непрерывной случайной величины квантилем поярдка α $q_\alpha(\xi)$ назовём:

$$P(\xi < q_\alpha(\xi)) = \alpha$$

Опр: квантилем порядка α случайнйо величины назовём

$$q_\alpha(\xi) = \inf\{a : P(\xi \geq a) \leq 1 - \alpha\}$$

то есть такое минимальное число a, для кототорого вероятность быть величины ξ принимать большие значения, чем α меньше $1 - \alpha$ или вероятность случайной величины принимать значения, меньшие, чем квантиль, равна α.

Если распределение непрерывно, то α квантиль однозначно задется уравнением:

$$F_\xi(q_\alpha(\xi)) = \alpha$$

где $F_\xi(x)$ - функция распределения случайной величины.

Опр: $\alpha = 0.25$-квантиль называется первым или нижним квартилем

Опр: $\alpha = 0.5$- квантиль называется медианой или вторым квартилем

Опр: $\alpha = 0.75$-квантиль называется третим или верхним квартилем

Опр: интервальным размахом называется разность между третьим и перывм квартилями, то есть $q_{0.75} - q_{0.25}$. Интервальный размах является характеристикой разброса распределения величины. Вместе, медиана и интервальный размах величины могут быть использованы вместо математического ожидания и дисперсии в случае распределений с большими выбросами, либо при невозможности вычислить последние.

2.5 Производящие и характеристические функции

2.5.1 Производящая функция вероятности

Пусть ξ - неотрицтальеная целочисленная случайная величина, то есть ξ принимает значения $0, 1, 2, \dots$ с вероятостями $p_k = P(\xi = k)$.

Опр: производящей функцией вероятности целочисленной неотрицательной случайной величины ξ назовём

$$\varphi_\xi(z) = Ez^\xi = \sum_{k=0}^{\infty} z^k p_k = z^0 p_0 + z^1 p_1 + z^2 p_2 + \dots$$

определенная для комплексных z таких, что $z \leq 1$, так как для подобных z ряд, определяющий производящую функцию, равномерно сходится, так как:

$$|\varphi(z)| = |\sum_{k=0}^{\infty} z^k p_k| \leq \sum_{k=0}^{\infty} |z|^k p_k \leq \sum_{k=0}^{\infty} p_k = 1$$

Производящая функция однозначно определяет распределение случайной величины (функция вероятности восстанавливается взятием производной):

$$p_k = p(\xi = k) = \frac{\varphi^{(k)}(0)}{k!}$$

Свойства производящей функции:

1. $\varphi_\xi(1) = \sum_{k=0}^{\infty} p_k = 1 \Rightarrow$ радиус сходимости не меньше единицы.

2. $\varphi'_\xi(1) = \sum_{k=1}^{\infty} kz^{k-1}p_k = \sum_{k=1}^{\infty} kp_k = E\xi$
3. $\varphi^{(n)}_\xi(z) = \sum_{k=n}^{\infty} k(k-1)\dots(k-n+1)z^{k-n}p_k$
4. $\varphi^{(2)}_\xi(z) = \sum_{k=2}^{\infty} k(k-1)z^{k-2}p_k|_{z=1} = \sum_{k=2}^{\infty} k^2p_k - \sum_{k=2}^{\infty} kp_k = E\xi^2 - E\xi$
5. Пусть неотрицательная целочисленная случайная величина ξ имеет момент n-ого порядка. Тогда:

$$\varphi^{(n)}_\xi(z)|_{z=1} = E(\xi(\xi-1)\dots(\xi-n+1))$$

6. $D\xi = E\xi^2 - (E\xi)^2 = E(\xi(\xi-1)) + E\xi - (E\xi)^2 = \varphi''_\xi(1) + \varphi'_\xi(1) - (\varphi'_\xi(1))^2$
7. Производящая функция одноначо определяет распределение целочисленной неотрицательной случайной величины, то есть производящие функции равны тогда и только тогда, когда ξ и η имеют одинаковые распределения.
8. $|\varphi_\xi(z)| \leq 1$ при $|z| \leq 1$
9. Для любых вещественныйх a и b в силу линейности математического ожидания:

$$\varphi_{a\xi+b}(z) = E(z^{a\xi+b}) = z^bE((z^a)^\xi) = z^b\varphi_\xi(z^a)$$

10. Если ξ и η дискретные независимые случайные величины со своими производящими функциями, то

$$\varphi_{\xi_1+\xi_2}(z) = E(z^{\xi_1+\xi_2}) = E(z^{\xi_1}z^{\xi_2}) = Ez^{\xi_1}Ez^{\xi_2} = \varphi_{\xi_1}(z)\varphi_{\xi_2}(z)$$

2.5.2 Производящие функции моментов

Моменты - ожидаемые значения случайной величины ξ. Первый момент - математическое ожидание - среднее значение случайной величины, дисперсия - мера разброса значений. Третий момент указывает на ассиметрию, а четвертый - насколько тяжелы хвосты распределения.

$$e^{t\xi} = 1 + t\xi + \frac{1}{2}t^2\xi^2 + \dots + \frac{1}{n!}t^n\xi^n + \dots$$

Опр: производящей функцией моментов случайной величины ξ назовём

$$\Psi_\xi(t) = Ee^{t\xi} = E(1 + t\xi + \frac{1}{2}t^2\xi^2 + \ldots + \frac{1}{n!}t^n\xi^n + \ldots) =$$

$$= 1 + tE\xi + \frac{1}{2}t^2E\xi^2 + \ldots + \frac{1}{n!}t^nE\xi^n + \ldots = \sum_{k=1}^{\infty}\frac{t^kE\xi^k}{k!}$$

Для непрерывной случайной величины ξ производящая функция моментов записывается следующим образом:

$$\Psi_\xi(t) = \int_x e^{tx}f_\xi(x)dx$$

Для дискретной случайной величины $P(\xi = x_k) = p_k, k = 1, 2, \ldots$:

$$\Psi_\xi(t) = Ee^{t\xi} = \sum_{k=0}^{\infty} e^{tx_k}p_k$$

Производящая функция моментов служит для выражения любого момента через нее путем взятия производной n-ой степени по t - вспомогательной прменной, чтобы использовать исчисление (производные) и обнулить значения, которые нас интересуют.

Конечно, моменты можно вычислить используя определение ожидаемых значений, но производящие функции моментов упрощают вычисление моментов. Чтобы производяшая функция моментов существовала, должно существовать ожидаемое значение $Ee^{t\xi}$, то есть интеграл должен сходиться (прочитать про сходимость интеграла).

При помощи производящей функции моментов можно находить моменты при помощи производных, а не интегралов.

Свойства производящих функций моментов:

1. $E\xi^k = \Psi_\xi^{(k)}(0)$
2. $\Psi_\xi(0) = 1$
3. если ξ_1 и ξ_2 независимые, то $\Psi_{\xi_1+\xi_2}(x) = \Psi_{\xi_1} \cdot \Psi_{\xi_2}$
4. Производящая функция моментов однозначно определяет распределение, в частности, елси обе случайные величины абсолютно непрерывны, то совпадение производящих функций велчет совпадение

плотностей. Если обе случайные велчиины дискретны, то совпадение влечет совпадение функций вероятности.

2.5.3 Характеристические функции

Характеристической функцией случайной величины ξ называется комплекснозначная функция $\phi(t) = E(e^{it\xi})$, определенная для всех действительных значений t.

Для непрерывной случайной величины ξ производящая функция моментов записывается следующим образом:

$$\phi_\xi(t) = \int_{-\infty}^{\infty} e^{itx} f_\xi(x) dx$$

Для дискретной случайной величины $P(\xi = x_k) = p_k, k = 1, 2, \ldots$:

$$\phi_\xi(t) = \sum_{k=1}^{\infty} e^{itx_k} p_k$$

Для случайной величины с произвольной функцией распределения $F_\xi(x)$ математическое оиждание $E(e^{it\xi})$ определяется с помощью функции интеграла Стилтьеса, т.е характеристическая функция определяется формулой:

$$\phi_\xi(t) = \int_{-\infty}^{\infty} e^{itx} dF_\xi(x)$$

Свойства характеристических функций:

1. Для целочисленных случайных величин $\phi_\xi(t) = \varphi_\xi(e^{it})$
2. $\phi_\xi(0) = E(e^0) = E(1) = 1$ - характеристическая функция в нуле равна единице
3. $|\phi_\xi(t)| \leq E|e^{it\xi}| = E(1) = 1$ - характеристическая функция всегда (!) ограничена
4. характетристическая функция всегда равномерно непрерывна
5. Для любых вещественных a и b:

$$\phi_{a\xi+b}(t) = Ee^{it(a\xi+b)} = Ee^{ita\xi} \cdot e^{itb} = e^{itb} Ee^{i(ta)\xi} = e^{itb} \phi_\xi(at)$$

6. Если ξ и η независимые случайные величины, то:

7. если ξ_1 и ξ_2 независимые, то $\phi_{\xi_1+\xi_2}(x) = \phi_{\xi_1} \cdot \phi_{\xi_2}$

8. если $\xi_1, \xi_2, \ldots, \xi_n$ - независимые случайные величины, то характеристическая функция суммы независимых случайных величин распадается в произведение характеристических функций слагаемых:

$$\phi_{\sum_{i=1}^{k} \xi_k}(t) = \prod_{i=1}^{k} \phi_{\xi_k}(t)$$

9. Лемма о вычислении моментов: пусть случайная величина ξ имеет абсолютный момент n-ого порядка. Тогда характеристическая функция случайной величины ξ дифференцируема n раз и при $0 \le k \le n$
$$E(\xi^k) = (-i)^k \phi_\xi^{(k)}(0)$$
$$\phi_\xi^{(k)}(0) = i^k E(\xi^k)$$

10. Математическое ожидание и дисперсия выражаются формулами:

$$E\xi = -i\phi'_\xi(0)$$

$$D\xi = E\xi^2 - (E\xi)^2 = (-i)^2\phi''_\xi(0) - (-i\phi'_\xi(0))^2 = -\phi''_\xi(0) + (\phi'_\xi(0))^2$$

Пусть существует момент порядка $k \in \mathbb{N}$ случайной величины ξ, т.е $E|\xi|^k < \infty$. Тогда характеристическая функция $\phi_\xi(t)$ в окрестности точки $t = 0$ разлагается в ряд Тэйлора:

$$\phi_\xi(t) = \phi_\xi(0) + \sum_{j=1}^{k} \frac{t^j}{j!}\phi_\xi^{(j)}(0) + o(|t^k|) = 1 + \sum_{j=1}^{k} \frac{i^j t^j}{j!} E\xi^j + o(|t^k|) =$$

$$= 1 + itE\xi - \frac{t^2}{2}E\xi^2 + \ldots + \frac{i^k t^k}{k!}E\xi^k + o(|t^k|)$$

Формулы, с помощью которых по характеристической функции восстанавливается распределение, в анализе называют формулами «обратного преобразования Фурье». Например, если модуль характеристической функции интегрируем на всей прямой, то у случайной величины есть плотность распределения, и она находится по формуле

$$f_\xi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi_\xi(t)dt$$

Формула обращения для целочисленной случайной величины ξ:

$$p_k = p(\xi = k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{itx}\phi(t)dt$$

Функция распределения однозначно определяется своей характеристической функцией. Утверждение справделиво, когда функция распределения отвечает целочисленной случайной величине или имеет плотность распределения.

3 Законы расределения случайных величин

3.1 Дискретные распределения

1. Распределение Бернулли

Будем говорить, что случайная величина имеет распределение бернулли, если вероятность 0 - $1 - p$, а 1 - p

$$E\xi = p, D\xi = E\xi^2 - (E\xi)^2 = 0^2 \cdot (1-p) + 1^2 \cdot p - p^2 = p - p^2$$

2. Биноминальное распределение $\xi \sim Bin(n, p)$

Пусть ξ - случайная величина, равная числу успехов в n испытаниях Бернулли и пусть вероятность успеха в каждом испытании p. Тогда ξ имеет биноминальное распределение с параметрами n и p, если её закон распределения:

$$P(\xi = k) = C_n^k p^k (1-p)^{n-k}, k = 0, 1, \ldots, n$$

Коэффициенты C_n^k нзываются биномиальными коэффициентами, так как они входят в формулу бинома Ньютона:

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}$$

Вычислим производящую функцию биноминального расппределения (последний шаг следует из формулы бинома Ньютона):

$$\varphi_\xi(z) = Ez^\xi = \sum_{k=0}^{\infty} z^k P(\xi = k) = \sum_{k=0}^{\infty} z^k C_n^k p^k (1-p)^{n-k} =$$

$$= \sum_{k=0}^{\infty} C_n^k (zp)^k (1-p)^{n-k} = (zp + 1 - p)^n$$

$$E\xi = \varphi'(z)|_{z=1} = np(zp + 1 - p)^{n-1} = np$$

$$E\xi^2 = \varphi''(z)|_{z=1} = (np(zp+1-p)^{n-1})'(z)|_{z=1} - \varphi'(z)|_{z=1} = np^2(n-1) - E\xi$$

$$D\xi = E\xi^2 - (E\xi)^2 = \varphi''_\xi(1) + \varphi'_\xi(1) - (\varphi'_\xi(1))^2 =$$

$$= np^2(n-1) - (np)^2 = n^2p^2 - np^2 - n^2p^2 + np = np - np^2 = np(1-p) = npq$$

Производящая функция моментов: $\phi_\xi(t) = (pe^t + q)^n$

Свойства биноминального распределения:

1. Пусть $\xi_1 \sim Bin(n,p)$ и $\xi_2 \sim Bin(n, 1-p)$. Тогда $p_{\xi_1}(k) = p_{\xi_2}(n-k)$
2. Пусть $\xi_1 \sim Bin(n_1,p)$ и $\xi_2 \sim Bin(n_2,p)$. Тогда $\xi_1 + \xi_2 \sim Bin(n_1+n_2, p)$

3. Распределение Пуассона $\xi \sim Poiss(\lambda)$

Это распределени является предельным для биномиального распределения, когда $n \to \infty$, а $\lambda = np$ остаётся постоянным. Говорят, что случайная величина ξ имеет распределение Пуассона с параметром $\lambda > 0$, если:

$$P(\xi = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k = 0, 1, 2, \dots$$

Вычислим производящую функцию распределения Пуассона:

$$\varphi_\xi(z) = Ez^\xi = \sum_{k=0}^{\infty} z^k P(\xi = k) = \sum_{k=0}^{\infty} z^k \frac{e^{-\lambda}\lambda^k}{k!} =$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}$$

$$E\xi = \varphi'(z)|_{z=1} = e^{\lambda(z-1)'}(z)|_{z=1} = \lambda e^{\lambda(z-1)}|_{z=1} = \lambda$$

$$D\xi = E\xi^2 - (E\xi)^2 = \varphi''_\xi(1) + \varphi'_\xi(1) - (\varphi'_\xi(1))^2 =$$

$$= \lambda^2 e^{\lambda(z-1)}|_{z=1} + \lambda e^{\lambda(z-1)}|_{z=1} - (\lambda e^{\lambda(z-1)}|_{z=1})^2 = \lambda$$

$$E\xi^2 = D\xi + (E\xi)^2 = \lambda + \lambda^2$$

а. Произодящая функция биномиальной случайной величины сходится к пуассоновской при $n \to \infty, \lambda = Const, np = \lambda, p \to 0$:

$$\varphi_\xi(z) = (zp + 1 - p)^n = (1 + p(z-1))^{\frac{\lambda}{p}} = (!) e^{\lambda(z-1)} \Rightarrow$$

Производящие функции равны, следовательно, равны и функции распределения.

б. Для факториальных моментов распределения справедлива общая формула:

$$E\xi^k = \lambda^k$$

в. Сумма независимых пуассоновских случайных величин также имеет распределение Пуассона: $\xi_1 + \xi_2 \sim Poiss(\lambda_1 + \lambda_2)$

г. С увеличением λ распределение Пуассона стремится к распределнию Гаусса со среднеквадратичным отклонением $\sigma = \sqrt{\lambda}$ и сдвигом λ. (Докажем позже)

4. Геометрическое распределение $\xi \sim Geom(p)$

Проводится бесконечная последовательность независимых одинаковых испытаний, в каждом из которых событие A наступает с вероятностью $p = p(A) > 0$. Пусть ξ - случайная величина,равная числу испытаний до момента первого наступления события A (количество испытаний до первого успеха), если имеет следующий закон распределения:

$$P(\xi = k) = (1-p)^k \cdot p, k = 0, 1, 2, \ldots$$

Это распределение называется геометрическим, так как вероятности $P(\xi = k)$ образуют геометрическую прогрессию.

Вероятность того, что событие A наступит не раньше момента m задается формулой:

$$P(\xi > m) = \sum_{k=m}^{\infty} (1-p)^k p = (1-p)^m p \frac{1}{1-(1-p)} = (1-p)^m$$

Производящая функция геометрического распределения определяется формулой:

$$\varphi_\xi(z) = Ez^\xi = \sum_{k=0}^{\infty} z^k P(\xi = k) = \sum_{k=0}^{\infty} z^k (1-p)^k \cdot p$$

$$= p \sum_{k=0}^{\infty} (z(1-p))^k = \frac{p}{1 - z(1-p)}$$

Для математического ожидания и дисперсии имеем:

$$E\xi = \frac{p(1-p)}{(1-z(1-p))^2} = \frac{1-p}{p}$$

$$D\xi = \varphi''_\xi(1) + \varphi'_\xi(1) - (\varphi'_\xi(1))^2 =$$

$$= \frac{2p(1-p)^2}{(1-z(1-p))^3}|_{z=1} + \frac{1-p}{p} - \left(\frac{1-p}{p}\right)^2 = \frac{1-p}{p^2}$$

Свойство потери памяти:

$P(\xi > t+s|\xi > t) = P(\xi > s)$ - не зависит от t, то есть количество прошлых неудач не влияет на количество будущиъ неудач.

Производящая функция моментов:

$$\varphi_\xi(t) = \frac{p}{1-qe^t}$$

3.1.1 Непрерывные распределения

Рассмотрим непрерывные распределения, а именно распределения, имеющие плотности.

5. Равномерное распределение $\xi \sim U[0;1]$

Будем говорить, что случайная величина равномерно распределена на отрезке $[a;b]$, если

$$f_\xi(x) = \frac{1}{b-a} I_{[a,b]}(x) = \begin{cases} \frac{1}{b-a}, a \le x \le b \\ 0, x \notin [a;b] \end{cases}$$

$$F_\xi(x) = \begin{cases} 0, \quad x < a \\ \frac{x-a}{b-a}, a \le x < b \\ 1, \quad x \ge b \end{cases}$$

Производящая функция моментов:

$$\Psi(t) = \int_{-\infty}^{\infty} e^{tx} f_\xi(x) dx = \frac{1}{b-a} \int_a^b e^{tx} dx = \frac{e^{tb} - e^{ita}}{t(b-a)}$$

Характеристическая функция равномерного распределения:

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f_\xi(x) dx = \frac{1}{b-a} \int_a^b e^{itx} dx = \frac{e^{itb} - e^{ita}}{it(b-a)}$$

Откуда находятся все интересующие моменты непрерывного равномерного распределения:

$$E\xi = \frac{a+b}{2}$$

$$D\xi = \frac{(b-a)^2}{12}$$

$$E\xi^2 = \frac{a^2+ab+b^2}{3}$$

Свойство:

$\xi \sim U[a;b], \eta = c\xi + d,$

Тогда: $\eta \sim U[ac+d, bc+d]$, $c > 0$ и наоборот, если $c > 0$

6. Экспоненциальное распределение $\xi \sim Exp(\lambda)$

Экспоненциальное (или показательное) распределение — абсолютно непрерывное распределение, моделирующее время между двумя последовательными свершениями одного и того же события.

Случайная величина ξ имеет экспоненциальное распределение с параметром $\lambda > 0$ (интенсивность или обратный коэффициент масштаба), если её плотность имеет вид:

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, x < 0 \\ 0, x \leq 0 \end{cases}$$

Функция распределения экспоненциальной величины:

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

Производящая функция моментов для эксоненциального распределения имеет вид:

$$\Psi_\xi(x) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = (1 - \frac{t}{\lambda})^{-1}$$

Характеристическая функция моментов для эксоненциального распределения имеет вид:

$$\phi_\xi(x) = \int_0^\infty e^{itx} \lambda e^{-\lambda x} dx = (1 - \frac{it}{\lambda})^{-1}$$

Откуда получаем все моменты:

$$E\xi^k = \frac{k!}{\lambda^k}$$

В частности:

$$E\xi = \frac{1}{\lambda}, E\xi^2 = \frac{2}{\lambda^2}, D\xi = E\xi^2 - (E\xi)^2 = \frac{1}{\lambda^2}$$

Свойства экспоненциального распределения:

1. Экспоненциальное распредление возникает как предельное для геометрического:

$$P(\xi \geq t) = \lim_{\Delta t \to 0} (1 - \lambda \Delta t)^{\frac{t}{\Delta t}} = e^{-\lambda t}$$

2. Свойство отсутствия памяти - если время ожидания распределено о экспоненциальному закону, то информация о том, что событие не наступило к данному моменту не улучшает шансы на его поступление в дальнейшем:

$$P(\xi > t + s | \xi > t) = \frac{P(\xi \geq t + s \cap \xi \geq t)}{P(\xi \geq t)} = \frac{P(\xi \geq t + s)}{P(\xi \geq t)} =$$

$$= \frac{1 - (1 - e^{-\lambda(t+s)})}{1 - (1 - e^{-\lambda t})} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P(\xi \geq s)$$

7. Нормальное распределение $\xi \sim N(\mu, \sigma^2)$

Будем говорить, что величина с μ - коэффициентом сдвига и $\sigma > 0$ - коэффициент масштаба, вещественный, строго положительный имеет нормальное распределение, если ее плотность вероятности имеет вид:

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Это расределение также называют гауссовским распределением.

Если $\mu = 0$ и $\sigma = 1$, то нормальное распределение с такими параметрами называется стандартным.

Функция распределения имеет вид:

$$F_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Функция стандартного нормального распределения имеет вид:

$$\Phi_\xi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} dt = \frac{1}{2}(1 + N(x))$$

где $N(x)$ - функция Лапласа.

$$N(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} dt$$

Характеристическая функция нормального распределения:

$$\phi_\xi(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$$

Производящая функция моментов нормального распределения:

$$\Psi_\xi(t) = e^{t\mu + \frac{\sigma^2 t^2}{2}}$$

Вычислим математическое ожидание и дисперсию нормального распределения:

$$E\xi = \mu, D\xi = \sigma^2$$

Таким образом, параметры μ и σ имеют следующий смысл - μ - математическое ожидание, σ - стандартное отклонение.

Используя функцию Лапласа можно выразить вероятность, что случайная величина ξ, распределенная по нормальному закону с параметрами a и σ лежит в заданных пределах. Справедлива следующая формула:

$$P(x_1 \leq \xi \leq x_2) = \frac{1}{2}(N(\frac{x_2 - \mu}{\sigma}) - N(\frac{x_1 - \mu}{\sigma}))$$

Легко убедиться, что $N(x)$ является функцией распределения величины $|\xi|$, где ξ - случайая величина со стандартным нормальным распределением. В силу этого свойств $N(x)$ $x \in (0, \infty)$ называют функцией отраженного нормального распределения. Для отрицательных x распределение следует полагать равным нулю.

Ведь если η - нормальная случайная величина, то величина $\xi = \sigma\eta + \mu$ будет иметь нормальное распределение с математическим ожиданием μ и стандартным отконением σ. Наоборот, если ξ - нормальная величина, то $\eta = \frac{\xi-\mu}{\sigma} \sim N(0,1)$, так как:

$$f_{\frac{\xi-\mu}{\sigma}} = \frac{\sigma}{\sigma\sqrt{2\pi}} e^{-\frac{(\sigma y+\mu-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

$$N(-x) = 1 - N(x) \Rightarrow \Phi_\xi(-x) = p(\xi < -x) = 1 - p(\xi < x) = 1 - \Phi_\xi(x)$$

Правило трех сигм:

Все значения нормально распределнной случайной величины лежат в интервала $(\mu-3\sigma; \mu+3\sigma)$. Приблизительно с 0.9973 вероятностью значение нормально распределенной случайной величины лежит в указанном интервале.

Доказательство:

Пусть $x_1 = -3\sigma + \mu$, $x_2 = 3\sigma + \mu$, тогда:

$$P(|\xi-\mu| < 3\sigma) = P(-3\sigma+\mu < \xi < 3\sigma+\mu) = P\left(\frac{-3\sigma+\mu-\mu}{\sigma} \leq \frac{\xi-\mu}{\sigma} < \frac{3\sigma+\mu-\mu}{\sigma}\right)$$

$$= \Phi_\xi(3) - \Phi_\xi(-3) = P(\xi < 3) - (1 - P(\xi < 3)) = 2P(\xi < 3) - 1 = 0.9973$$

$$P(|\xi-\mu| < 3\sigma) = 2P(\xi < 2) - 1 = 0.9545$$

$$P(|\xi-\mu| < 3\sigma) = 2P(\xi < 1) - 1 = 0.682689 \blacksquare$$

Двумерное нормальное распределение.

Пусть ξ и η - независимые нормально распределенные случайные величины. $\xi_1 \sim N(\mu_1, \sigma_1^2)$,$\xi_2 \sim N(\mu_2, \sigma_2^2)$. Тогда $\xi_1+\xi_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

Совместная плотность распределения равна:

$$f_{\xi_1,\xi_2}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} \cdot e^{(-\frac{(x-\mu_1)^2}{2\sigma_1} - \frac{(x-\mu_2)^2}{2\sigma_2})}$$

Пусть $\xi_1 \sim N(\mu_1, \sigma_1^2)$,$\xi_2 \sim N(\mu_2, \sigma_2^2)$, $\rho = \rho(\xi_1, \xi_2)$.

Тогда:

$$\xi_1 + \xi_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$$

3.2 Виды сходимости

Напомню, что случайная величина есть (измеримая) функция из некоторого абстрактного множества Ω в множестве действительных чисел. Последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ есть, тем самым, последовательность функций, определенных на одном и том же пространстве элементарных исходов Ω. Существуют разные виды сходимости последовательности функций. Всякий раз давать определение какой-либо сходимости мы будем, опираясь на сходимость числовых последовательностей как на уже известное основное понятие.

В частности, при каждом новом $\omega \in \Omega$ мы имеем новую числовую последовательность $\xi_1(\omega), \xi_2(\omega), \xi_3(\omega), \ldots$. Поэтому можно говорить о знакомой из математического анализа поточечной сходимости последовательностей функций: о сходимости "почти всюду которую в теории вероятностей называют "почти наверное".

Опр: говорят, что последовательность случайных величин ξ_n сходится **почти наверное** (почти всюду, с вероятностью 1) к случайной величине ξ, если

$$\xi_n \xrightarrow[n \longrightarrow \infty]{\text{п.н.}} \xi$$

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

Так как случайная величина - это функция, то в данном случае речь идет о поточечной сходимости.

Опр: будем говорить, что последовательность случайных величин ξ_n сходится по вероятности (p) к ξ, если для любого $\epsilon > 0$:

$$\xi_n \xrightarrow[n \longrightarrow \infty]{\text{p}} \xi$$

$$P\{|\xi_n - \xi| > \epsilon\} \xrightarrow[n \longrightarrow \infty]{} 0$$

$$P\{|\xi_n - \xi| \leq \epsilon\} \xrightarrow[n \longrightarrow \infty]{} 1$$

Опр: будем говорить, что последовательность случайных величин ξ_n сходится **по распределению** к случайной величине ξ:

$$\xi_n \xrightarrow[n \longrightarrow \infty]{\text{в}} \xi$$

если для любой непрерывной и ограниченной функции g(x):

$$Eg(\xi_n) \xrightarrow[n \longrightarrow \infty]{} Eg(\xi)$$

Опр: будем говорить, что последовательность случайных величин ξ_n **слабо сходится по распределению** к случайной величине ξ:

$$\xi_n \xrightarrow[n \longrightarrow \infty]{\text{F}} \xi$$

если для любого x такого, что функция распредлеения F_ξ непрерывна в точке x, имеет сходимость:

$$F_{\xi_n}(x) \xrightarrow[n \longrightarrow \infty]{} F_\xi(x)$$

3.3 Некоторые неравенства о случайных величинах

1. Неравенство Коши-Буняковского-Шварца

$$E|\xi_1\xi_2| \leq (E|\xi_1|^2)^{\frac{1}{2}} \cdot (E|\xi_2|^2)^{\frac{1}{2}}$$

2. Неравенство Гёльдера

$$E|\xi_1\xi_2| \leq (E|\xi_1|^r)^{\frac{1}{r}} \cdot (E|\xi_2|^l)^{\frac{1}{l}}$$

$$\frac{1}{r} + \frac{1}{l} = 1, r > 1$$

$$r, l > 0$$

3. Неравенство Миньковского

$$(E|\xi_1 + \xi_2|^r)^{\frac{1}{r}} \leq (E|\xi_1|^r)^{\frac{1}{r}} + (E|\xi_2|^r)^{\frac{1}{r}}$$

4. Неравенство Маркова

Неравенство Маркова в теории вероятностей даёт оценку вероятности, что случайная величина превзойдёт по модулю фиксированную положительную константу, в терминах её математического ожидания. Хотя получаемая оценка обычно груба, она позволяет получить определённое представление о распределении, когда последнее не известно явным образом.

Пусть неотрицательная случайная величина ξ определена на вероятностом пространстве и её математическое ожидание конечно. Тогда

$$\forall \epsilon > 0 : P(\xi \geq \epsilon) \leq \frac{E\xi}{\epsilon}$$

4.1 Следствие из неравенства Маркова:

$$P(|\xi| \geq \epsilon) \leq \frac{E|\xi|}{\epsilon}$$

5. Неравенство Чебышёва

Неравенство Чебышёва в теории вероятностей утверждает, что случайная величина в основном принимает значения, близкие к своему среднему. А более точно, оно даёт оценку вероятности того, что случайная величина примет значение, далёкое от своего среднего.

Пусть случайная веичина ξ определена на вероятностом пространстве, а ее математическое ожидание $E\xi$ и дисперсия $D\xi$ конечны. Тогда:

$$\forall \epsilon > 0 : P(|\xi - E\xi| \geq \epsilon) \leq \frac{D\xi}{\epsilon^2}$$

Если $\epsilon = k\sigma$, где σ - стандартное отклонение, а $k > 0$, то получаем, что:

$$P(|\xi - E\xi| \geq k\sigma) \leq \frac{1}{k^2}$$

Пример: вероятность того, что случайная величина будет отклоняться от своего среднего на расстояние больше 3σ записывается следующим неравенством:

$$P(|\xi - E\xi| \geq 3\sigma(\epsilon)) \leq \frac{\sigma^2}{3^2\sigma^2} \leq \frac{1}{9}$$

3.4 Закон больших чисел

Пусть проводится большое количество независимых одинаковых экспериментов, в каждом из которых наблюдается случайная величина одной и той же природы. С математической точки зрения это означает, что наблюдается последовательность независимых случайных величин с одинаковыми распределениями. Предполагается, что существует математическое ожидание отдельно взятой величины. Обозначим его за μ. Заметим, что математические ожидания у одинаково распределенных случайных величин одинаковы. Тогда среднее арифметическое всех значений случайных величин, полученных в результате экспериментов, прближается при возрастании числа экспериментов к неслучайному числу μ.

Теорема: Закон больших чисел

Пусть $\{\xi_j\}_{j=1}^{\infty}$ - независимые, одинаково распределенные случайные величины с математическим ожиданием $\mu = E\xi_j$ и конечными дисперсиями $D\xi_j = \sigma^2 < \infty$.

Тогда величина $\bar{\xi_n} = \frac{1}{n}\sum_{j=1}^{n} \xi_i$, равная среднему арифметическому первых n величин последовательности ξ_n, сходится при $n \to \infty$ по вероятности и с вероятностью единица к математическому ожиданию μ.

$$\frac{S_n}{n} \xrightarrow[n\longrightarrow\infty]{p} \mu$$

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \mu\right| > \epsilon\right\} \xrightarrow[n\longrightarrow\infty]{} 0$$

Суть: хотя отдельные независимые случайные величины могут принимать значения, далекие от своих мат.ожиданий, среднее арифметическое достаточно большого числа случайных величин с большой вероятностью принимает значение, близкое к определенному постоянному числу. Т.е. разброс среднего арифметического будет мал.

Интерпретируя данный результат получаем, что слабый закон утверждает, что для любых ненулевых указанных границ, независимо от того, насколько они малы, при достаточно большой выборке вероятность того, что среднее значение выборки будет близко к математическому ожиданию, очень высока в пределах этих границ.

Как говорилось ранее, слабый закон применим в случае независимых одинаково распределенных случайных величин, имеющих математическое ожидание. Однако он может применяться и в некоторых других случаях. Например, дисперсия может быть разной для каждой случайной величины в выборке, а математическое ожидание оставаться константой. Если дисперсии ограничены, то закон также применим, как показал Чебышёв ещё в 1867 году. Доказательство Чебышёва работает до тех пор, пока дисперсия среднего числа первых n значений не стремится к нулю при $n \to \infty$

3.5 Центральная предельная теоерма Ляпунова

Лемма 1. Пусть есть некоторые случайные величины $\{\xi_n\}$ с функцией распределения $F_{\xi n}(x)$ и характеристической функцией $\phi_{\xi n}(t)$ и есть случайная величина ξ с функцией распределения $F_\xi(x)$ и характеристической функцией $\phi_\xi(t)$. Тогда равносильные следующие утверждения:

1. В каждой точке непрерывности функции $F_\xi(x)$ имеет место сходимость:

$$F_{\xi n}(x) \xrightarrow[n\longrightarrow\infty]{} F_\xi(x)$$

$$\phi_{\xi n}(t) \xrightarrow[n\longrightarrow\infty]{} \phi_\xi(t), \forall t \in \mathbb{R}^1$$

Лемма 2. Если функция распределения $F_\xi(x)$ непрерывна на всей оси, то поточечная сходимость $F_{\xi n}(x) \to F_\xi(x)$ для любого $x \in \mathbb{R}^1$ эквивалентно равномерной сходимости при $n \to \infty$:

$$sup|F_{\xi n}(x) - F_\xi(x)| \xrightarrow[n\longrightarrow\infty]{} 0$$

Теорема: Центральная предельная теорема Ляпунова

Пусть $\{\xi_j\}_{j=1}^{\infty}$ - попарно независимые, одинаково распределенные случайные величины имеют конеченые математическое ожидание $E\xi_j = \mu < +\infty$ и дисперсию $D\xi_j = \sigma^2 < +\infty$. $S_n = \sum_{i=1}^{n} \xi_i$

Обозначим за:

$$\eta_n = \frac{\frac{1}{n}\sum_{j=1}^{n}\xi_j - \mu \cdot n}{\sigma\sqrt{n}}$$

Пусть $F_{\eta n}(x) = P(\eta < x)$ - функция распределения введенной случайной величины η. Тогда при $n \to \infty$для любого $x \in R$:

$$sup|F_{\xi n}(x) - \Phi(x)| \xrightarrow[n\longrightarrow\infty]{} 0$$

$$F_{\eta_n}(x) \xrightarrow[n\longrightarrow\infty]{} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

$$\eta_n = \frac{\frac{1}{n}\sum_{j=1}^{n}\xi_j - \mu \cdot n}{\sigma\sqrt{n}} \xrightarrow[n\longrightarrow\infty]{\text{по распределению}} N(0,1)$$

$\Phi(x)$ - функция распределения стандартного нормального распределения, к ней осуществляется поточечная сходимость.

Суть: при выполнении определенных условий, например, случайные величины одинаково распределены и попарно независимы, среднее арифметическое большого числа случайных величин имеет распределение близкое к нормальному.

Мы можем переписать результа центральной предельной теоремы в следующем виде:

$$\sqrt{n} \cdot \frac{\overline{X} - \mu}{\sigma} \xrightarrow[n\longrightarrow\infty]{\text{по распределению}} N(0,1)$$

где $\overline{X} = \frac{1}{n}\sum_{i=1}^{n}\xi_i$ - выборочное среднее (уже совсем скоро дадим точное формальное определение).

Неформально говоря, классическая цетральная предельная теорема утверждает, что сумма n независимых одинаково распределенных случайных величин имеет распределение, близкое к $\mathbb{N}(n\mu, n\sigma^2)$, что эквивалентно:

$$\overline{X} \sim \mathbb{N}(\mu, \frac{\sigma^2}{n}), \sum_{i=1}^{n} \xi_i \sim \mathbb{N}(\mu \cdot n, \sigma^2 \cdot n)$$

$$P(s_1 \leq S_n \leq s_2) = P\left(\frac{s_1 - n \cdot \mu}{\sigma\sqrt{n}} \leq \frac{S_n - n \cdot \mu}{\sigma\sqrt{n}} < \frac{s_2 - n \cdot \mu}{\sigma\sqrt{n}}\right) =$$

$$= \Phi\left(\frac{s_2 - n \cdot \mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{s_1 - n \cdot \mu}{\sigma\sqrt{n}}\right) = \frac{1}{2}\left(N\left(\frac{s_2 - n \cdot \mu}{\sigma\sqrt{n}}\right) - N\left(\frac{s_1 - n \cdot \mu}{\sigma\sqrt{n}}\right)\right)$$