Теория вероятностей и математическая статистика

доц. Флегель Александр Валерьевич

flegel@cs.vsu.ru

2014

Закон больших чисел

Экспериментальный факт: ***в длинной серии опытов частота появления события*** A ***сближается с определенным числом*** $P(A)$

Получим оценки распределений случайных величин.

Теорема (1)

Пусть $X = X(\omega) \geq 0 \ \forall \omega \in \Omega$. *Если* $M(X)$ *существует, то* $\forall \varepsilon > 0$

$$P(X \geq \varepsilon) \leq \frac{M(X)}{\varepsilon}$$

Доказательство:
Докажем для случая одномерного непрерывного вероятностного пространства.
Пусть $\Omega = \{u: \ -\infty < u < \infty\}$; $X(u) = x$; $f(x)$ – плотность распределения X.

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Пусть $\Omega_\varepsilon = \{u: \ X(u) \geq \varepsilon\} \subseteq \Omega$. Введем случайную величину Y:

$$Y = \begin{cases} 0, & u \in \Omega \backslash \Omega_\varepsilon \\ \varepsilon, & u \in \Omega_\varepsilon \end{cases}$$

$$\forall u \in \Omega: \quad X \geq Y$$

Умножим последнее неравенство на $f(x)$ и проинтегрируем.

$$M(X) \geq \varepsilon P(\Omega_\varepsilon) = \varepsilon P(X \geq \varepsilon). \quad \square$$

Теорема (2: неравенство Чебышева)

Если случайная величина X имеет дисперсию, то $\forall\varepsilon > 0$

$$P(|X - M(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}.$$

Доказательство:

$$\forall\omega \in \Omega: \quad Y = (X - M(X))^2 \geq 0 \quad \text{и} \quad \exists \ M(Y) = D(X).$$

Следовательно (см. предыдущую теорему):

$$P(|X - M(X)| \geq \varepsilon) = P(Y \geq \varepsilon^2) \leq \frac{M(Y)}{\varepsilon^2} = \frac{D(X)}{\varepsilon^2}.$$

□

Неравенство Чебышева позволяет оценивать вероятности отклонений значений случайной величины от своего математического ожидания.

Пример 1 (применение неравенства Чебышева)

Нужно оценить долю (M/N) бракованных изделий в партии, содержащей N изделий.

Отберем случайно (без возвращений) n изделий. Пусть среди отобранных X бракованных. X подчиняется гипергеометрическому распределению:

$$P(X = m) = \frac{C_M^m C_{N_M}^{n-m}}{C_N^n}$$

1. Математическое ожидание:

$$X = X_1 + X_2 + \cdots + X_n, \quad X_k = I_{A_k}, \quad (k = 1, 2, \ldots, n)$$

$A_k = \{$ k-тое изделие – бракованное $\}$

$$M(X_k) = M/N, \quad M(X) = nM/N.$$

Пример 1. Продолжение

2. Дисперсия $D(X) = M(X^2) - [M(X)]^2$:

$$
\begin{aligned}
M(X^2) &= M[(X_1 + \cdots X_n)^2] = M(X) + \sum_{k \neq j} M(X_k X_j) \\
&= \frac{nM}{N} + n(n-1)P(X_k = X_j = 1)
\end{aligned}
$$

$$
P(X_k = X_j = 1) = \frac{M(M-1)}{N(N-1)}
$$

$$
\begin{aligned}
D(X) &= n(n-1)\frac{M(M-1)}{N(N-1)} + \frac{nM}{N} - \frac{n^2 M^2}{N^2} \\
&= n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}
\end{aligned}
$$

Пример 1. Продолжение

Доля (M/N) бракованных изделий в партии, содержащей N изделий.

$$P\left(\left|\frac{X}{n}-\frac{M}{N}\right|\geq\Delta\right)\leq\frac{1}{n\Delta^2}\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$$

X/n – доля бракованных изделий в выборке из n изделий

Δ – некоторое положительное число

Пример 2 (применение неравенства Чебышева)

Оценка ошибки приближенного значения измеряемой величины

Пусть проводится n независимых измерений некоторой величины a.
$\delta_1, \delta_2, \ldots, \delta_n$ – ошибки измерения (случайные величины).
Предположим $M(\delta_k) = 0$ (отсутствие систематической ошибки).
Пусть $D(\delta_k) = b^2$.

Ошибка в определении числа a равна $X_n = (\delta_1 + \delta_2 + \cdots + \delta_n)/n$

$$M(X_n) = 0, \qquad D(X_n) = (D\delta_1 + \cdots + D\delta_n)/n^2$$

Предположим, что нам нужно, чтобы

$$P(|X_n| < \Delta) > 0.99 \quad \Leftrightarrow \quad P(|X_n| \geq \Delta) \leq 0.01$$

По неравенству Чебышева

$$P(|X_n| \geq \Delta) \leq \frac{D(X_n)}{\Delta^2} = \frac{b^2}{(n\Delta)^2} \quad \Rightarrow \quad n \geq 100\frac{b^2}{\Delta^2}.$$

Теорема (3)

Если случайные величины $X_1, X_2, \ldots, X_n, \ldots$ *попарно независимы и*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} DX_k = 0,$$

то $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left| \frac{X_1 + \cdots + X_n}{n} - \frac{MX_1 + \cdots + MX_n}{n} \right| < \varepsilon \right) = 1.$$

Доказывается введением случайной величины $Y_n = (X_1 + \cdots + X_n)/n$ и использованием неравенства Чебышева: $P(|Y_n - MY_n| \geq \varepsilon) \leq DY_n/\varepsilon^2$.

Замечание: вместо попарной независимости величин X_k можно потребовать некоррелированности: $\mathrm{cov}(X_k, X_j) = 0$ $(k \neq j)$.

Если выполнено утверждение теоремы (3), то говорят, что к случайным величинам применим **закон больших чисел**.

Частные случаи теоремы (3)

Теорема (4: теорема Чебышева)

Если случайные величины $X_1, X_2, \dots, X_n, \dots$ *попарно независимы и*

$$DX_k \le C, \quad C = \mathrm{const}$$

то $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \dots + X_n}{n} - \frac{MX_1 + \dots + MX_n}{n}\right| < \varepsilon\right) = 1.$$

Теорема (5)

Если случайные величины $X_1, X_2, \ldots, X_n, \ldots$ одинаково распределены, попарно независимы и имеют конечные дисперсии, то $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \cdots + X_n}{n} - a\right| < \varepsilon\right) = 1,$$

где $a = MX_k$.

Математической моделью последовательности из n измерений случайной величины a является случайный вектор $(X_1, X_2, \ldots, X_n)$, где X_k независимы и одинаково распределены.

Среднее арифметическое при больших n мало отличается от измеряемой величины с вероятностью, близкой к 1.

Теорема (6: теорема Бернулли)

Пусть μ_n – число успехов в n испытаниях Бернулли и p – вероятность успеха в отдельном испытании. Тогда $\forall \varepsilon > 0$

$$\lim_{n \to \infty} P\left(\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right) = 1.$$

Согласно теореме Бернулли частота μ_n/n наступления события A сближается с вероятностью p. Этот же факт установлен экспериментально.

Производящие функции

Определение

Если случайная величина ξ принимает только целые неотрицательные значения ($\xi = k,\ k = 1, 2, \ldots$), т.е.

$$\sum_{k=0}^{\infty} P(\xi = k) = 1,$$

то **производящей функцией** распределения ξ называется функция

$$\varphi_\xi(x) = M[x^\xi] = \sum_{k=0}^{\infty} x^k p_k, \qquad (1)$$

где $p_k = P(\xi = k)$, x – действительная или комплексная переменная ($|x| \leq 1$).

Теорема (1)

Пусть $\varphi_\xi(x)$ – производящая функция распределения ξ. Тогда:

1. $\varphi_\xi(x)$ *определена в каждой точке отрезка* $[-1, 1]$.
2. $\varphi_\xi(x) = 1$.
3. *Соответствие (1) между множеством* $\varphi_\xi(x)$ *и множеством распределений* $\{p_k\}$ *является взаимно однозначным.*

Используя коэффициенты ряда Тэйлора, можно явно указать распределение $\{p_k\}$:

$$p_k = \frac{1}{k!}\varphi_\xi^{(k)}(0).$$

Пример 1.

Пусть случайная величина имеет биномиальное распределение

$$P(x = m) = C_n^m p^m q^{n-m}, \quad m = 0, 1, \ldots, n.$$

По формуле (1)

$$\varphi_\xi(x) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} x^k = (px + q)^n.$$

По производящей функции легко определять моменты случайной величины.

Определение:

Моментом порядка k случайной величины ξ называется число $M(\xi^k)$.
Центральным моментом порядка k называется число $M[(\xi - M\xi)^k]$.
Факториальный момент порядка k – мат. ожидание величины $\xi^{[k]} = \xi(\xi-1)\cdots(\xi-k+1)$

$$M\xi = M\xi^{[1]}, \quad D\xi = M\xi^{[2]} + M\xi - (M\xi)^2$$

Теорема (2)

Если конечен k-й факториальный момент, то существует левосторонняя производная $\varphi_\xi^{(k)}(1)$ и

$$M\xi^{[k]} = \varphi_\xi^{(k)}(1), \qquad M\xi = \varphi'_\xi(1).$$

Пример 2.

Пусть случайная величина ξ имеет биномиальное распределение. Найдем $M\xi$, $D\xi$.

$$\varphi_\xi(x) = (px+q)^n,$$
$$\varphi'_\xi(x) = np(px+q)^{n-1},$$
$$\varphi''_\xi(x) = n(n-1)p^2(px+q)^{n-2},$$

$$M\xi = \varphi'_\xi(1) = np(p+q)^{n-1} = np,$$
$$M[\xi(\xi-1)] = \varphi''_\xi(1) = n(n-1)p^2(p+q)^{n-2} = n(n-1)p^2,$$
$$D\xi = n(n-1)p^2 + np - (np)^2 = np(1-p) = npq.$$

k-й факториальный момент биномиального распределения:

$$M\xi^{[k]} = n^{[k]}p^k, \quad n^{[k]} = n(n-1)\cdots(n-k+1).$$

Теорема (3)

Если случайные величины $\xi_1, \dots, \xi_n$ *независимы, то*

$$\varphi_X(x) = \varphi_{\xi_1}(x) \cdots \varphi_{\xi_n}(x) \qquad (X = \xi_1 + \dots + \xi_n).$$

Теорема (4)

Пусть целочисленные величины $\nu, \xi_1, \xi_2, \dots, \xi_n$ *независимы при любом* $n = 1, 2, \dots$; *и* ξ_k *одинаково распределены. Положим*

$$X_\nu = \xi_1 + \xi_2 + \dots + \xi_\nu, \quad X_0 = 0.$$

Тогда

$$\varphi_{X_\nu}(x) = \varphi_\nu[\varphi_{\xi_1}(x)].$$

Доказательство теоремы 4:

Вероятность события $\{X_\nu = m\}$ представим в виде

$$P(X_\nu = m) = \sum_{k=0}^{\infty} P(\nu = k, X_\nu = m).$$

1) $\{\nu = k, X_\nu = m\} = \{\nu = k, X_k = m\}$.
2) События $\{\nu = k\}$ и $\{X_k = m\}$ независимы:

$P(\nu = k, X_\nu = m) = P(\nu = k, X_k = m) = P(\nu = k)P(\xi_1 + \dots + \xi_k = m).$

$$P(X_\nu = m) = \sum_{k=0}^{\infty} P(\nu = k)P(\xi_1 + \dots + \xi_k = m).$$

Умножим обе части этого равенства на x^m и просуммируем по m:

$$\varphi_{X_\nu}(x) = \sum_{k=0}^{\infty} P(\nu = k)\left[\sum_{m=0}^{\infty} x^m P(\xi_1 + \dots + \xi_k = m)\right].$$

Доказательство теоремы 4 (продолжение):

$$\sum_{m=0}^{\infty} x^m P(\xi_1 + \dots + \xi_k = m) = \varphi_{\xi_1+\dots+\xi_k}(x) = [\varphi_{\xi_1}(x)]^k \,.$$

Следовательно,

$$\varphi_{X_\nu}(x) = \sum_{k=0}^{\infty} P(\nu = k)\,[\varphi_{\xi_1}(x)]^k = \varphi_\nu[\varphi_{\xi_1}(x)].$$

□

Теорема (5)

Пусть при любом фиксированном n ($n = 1, 2, \ldots$) последовательность $\{p_k(n)\}$ является распределением вероятностей, т.е.

$$p_k(n) \geq 0, \qquad \sum_{k=0}^{\infty} p_k(n) = 1.$$

Для того чтобы $\forall k$

$$\lim_{n\to\infty} p_k(n) = p_k, \qquad \sum_{k=0}^{\infty} p_k = 1,$$

необходимо и достаточно, чтобы $\forall x \in [0, 1)$

$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x),$$

где $\quad \varphi_n(x) = \sum_{k=0}^{\infty} p_k(n) x^k, \quad \varphi_n = \sum_{k=0}^{\infty} p_k x^k, \quad \varphi(1) = 1.$

Пример.

Пусть μ_n – число успехов в n испытаниях Бернулли и p_n – вероятность успеха в одном испытании. Будем предполагать, что $\exists \lim\limits_{n\to\infty} np_n = \lambda$. Воспользуемся теоремой 5 для вычисления $\lim\limits_{n\to\infty} P(\mu_n = m)$.

Пусть $\lambda_n = np_n$.

$$\varphi_{\mu_n}(x) = (p_n x + 1 - p_n)^n = (1 + p_n(x-1))^n = \left(1 + \frac{\lambda_n}{n}(x-1)\right)^n$$

$$\lim_{n\to\infty}\left(1+\frac{t}{n}\right)^n = e^t \quad \Rightarrow \quad \lim_{n\to\infty} \varphi_{\mu_n}(x) = e^{\lambda(x-1)} = \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} x^m.$$

$$\lim_{n\to\infty} P(\mu_n = m) = \frac{\lambda^m}{m!} e^{-\lambda}.$$

Характеристические функции

Производящие функции определены для целочисленных случайных величин. Для исследования распределений произвольных случайных величин вводятся **характеристические функции**.

Определение:

Характеристической функцией действительной случайной величины ξ называется

$$f_\xi(t) = Me^{it\xi},$$

где t – действительное число, $-\infty < t < \infty$.

Если ξ – дискретная, то

$$f_\xi(t) = \sum_{k=1}^{\infty} e^{itx_k} P(\xi = x_k).$$

Если ξ – абсолютно непрерывная величина, то

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

где $p_\xi(x)$ – плотность распределения ξ.

Свойства характеристической функции $f_\xi(t)$:

1. $f_\xi(t)$ определена $\forall t \in (-\infty, \infty)$.
2. $f_\xi(0) = 1$, $|f_\xi(t)| \leq 1$.
3. Если $Y = aX + b$, где a и b – постоянные, то $f_Y(t) = e^{itb} f_X(at)$.
4. Соответствие $f_\xi(t) \leftrightarrow$ множество функций распределения величины ξ является взаимно однозначным.