

1. Введение в математическую статистику

Мат. статистика - теория статистических решений. Основная задача - по экспериментальным данным высказать суждение о природе случайного явления (Оптимальное стат. решение).

Пример 1. В городе N жителей, среди них M заболевших. В результате осмотра n жителей выявлено m заболевших. Как можно оценить M ?

1.1. Сходимости случайных величин и векторов

Пусть $\xi, \{\xi_n\}$ - случайные векторы размерности m . Тогда

1. $\xi_n \xrightarrow{a.s.} \xi$, если $P(\lim_{n \rightarrow \infty} \xi_n = \xi) = 1$
2. $\xi_n \rightarrow^p \xi$, если $\forall \varepsilon > 0 P(|\xi_n - \xi| > \varepsilon) \rightarrow 0, n \rightarrow \infty$
3. $\xi_n \rightarrow^d \xi$, если $\forall f(x) : R^n \rightarrow R$ - огр. и непр. выполнено: $Ef(\xi_n) \rightarrow Ef(\xi), n \rightarrow \infty$

Теорема 1. Пусть $\xi, \{\xi_n\}_{n=1}^\infty$ - СВ. Тогда $\xi_n \rightarrow^d \xi \Leftrightarrow F_{\xi_n} \rightarrow^w F_\xi \Leftrightarrow F_{\xi_n} \Rightarrow F_\xi$

Теорема 2. Пусть $\xi, \{\xi_n\}_{n=1}^\infty$ - случайные векторы размерности m . Пусть $F_\xi(x)$ непрерывна. Тогда $\xi_n \rightarrow^d \xi \Leftrightarrow \forall x \in R^m F_{\xi_n} \rightarrow F_\xi, n \rightarrow \infty$

Теорема 3 <О соотношении видов сходимости>. Пусть $\xi, \{\xi_n\}_{n=1}^\infty$ - случайные векторы размерности m . Тогда

1. $\xi_n \rightarrow^{a.s.} \xi \Rightarrow \xi_n \rightarrow^p \xi$
2. $\xi_n \rightarrow^p \xi \Rightarrow \xi_n \rightarrow^d \xi$

Доказательство.

1). Пусть $\xi_n \rightarrow^{a.s.} \xi$

$$\Leftrightarrow \forall j = 1 \dots m \xi_n^{(j)} \rightarrow^{a.s.} \xi^{(j)} \Rightarrow \forall j = 1 \dots m \xi_n^{(j)} \rightarrow^p \xi^{(j)} \Leftrightarrow \xi_n \rightarrow^p \xi$$

2). Док-во полностью аналогично 1-мерному случаю (для СВ). □

Теорема 4 <без доказательства>. Пусть $\xi, \{\xi_n\}_{n=1}^\infty$ - случайные векторы размерности m . Если $\xi_n \rightarrow^p \xi$, то существует такая $\{\xi_n k\}$, что $\xi_n k \rightarrow^{a.s.} \xi$

Теорема 5 <ЗБЧ>. Пусть $\{\xi_n\}_{n=1}^\infty$ - непрер. СВ с условием $D\xi_n \leq C$. Положим $S_n = \xi_1 + \dots + \xi_n$. Тогда $\frac{S_n - ES_n}{n} \rightarrow^p 0$

Теорема 6 <УЗБЧ>.

Пусть $\{\xi_n\}_{n=1}^\infty$ - нез. о.р. СВ с ограниченной дисперсией. Обозначим S_n (аналогично). Тогда $\frac{S_n - ES_n}{n} \rightarrow^{a.s.} 0$

Теорема 7 <Центрально-предельная>.

Пусть $\{\xi_n\}_{n=1}^\infty$ - непрер. СВ с условием $0 < D\xi_n = \sigma^2 < +\infty$. обозначим S_n аналогично, $E\xi_n = a$. Тогда $\frac{S_n - ES_n}{\sqrt{DS_n}} \rightarrow^d N(0, 1)$

Теорема 8 <о наследовании сх-ти>.

Пусть $\xi, \{\xi_n\}_{n=1}^\infty$ - случайные векторы размерности m .

1. Если $\xi_n \rightarrow^{a.s.} \xi$ и $h(x) : R^m \rightarrow R^l$ такова, что h непрерывна почти всюду относительно распределения ξ . Т.е. $\exists B \in B(R^m) : h(x)$ непрерывна на B и $P(\xi \in B) = 1$. Тогда $h(\xi_n) \rightarrow^{a.s.} h(\xi)$
2. Если $\xi_n \rightarrow^p \xi$ и $h(x) : R^m \rightarrow R^l$ такова, что h непрерывна почти всюду относительно распределения ξ . Тогда $h(\xi_n) \rightarrow^p h(\xi)$
3. Если $\xi_n \rightarrow^d \xi$ и $h(x) : R^m \rightarrow R^l$ - непрерывна. Тогда $h(\xi_n) \rightarrow^d h(\xi)$

Доказательство.

1. $1 \geq P(\lim_{n \rightarrow \infty} h(\xi_n) = h(\xi)) = P(\lim_{n \rightarrow \infty} h(\xi_n) = h(\xi), \xi \in B) \geq P(\lim_{n \rightarrow \infty} \xi_n = \xi, \xi \in B) = 1 \Rightarrow P(\lim_{n \rightarrow \infty} h(\xi_n) = h(\xi)) = 1$
2. Пусть $h(\xi_n) \not\rightarrow^p h(\xi)$. Тогда $\exists \varepsilon_0 > 0 : \exists$ подпослед. $\xi_{nk}, \exists \delta_0 > 0 : P(|h(\xi_{nk}) - h(\xi)| > \varepsilon_0) \geq \delta_0 \forall k$. Но $\xi_{nk} \rightarrow^p \xi \Rightarrow$ есть еще подпослед. $\xi_{nk_s} : \xi_{nk_s} \xrightarrow{a.s.} \xi, s \rightarrow \infty$
3. Возьмем $f(x) : R^m \rightarrow R^m$ -огр, непрер. ф-я. $Ef(h(\xi_n)) \rightarrow^? Ef(h(\xi))$. Но $f(h(x))$ - непрерывная ограниченная в R^n и $\xi_n \rightarrow^d \xi$. Отсюда $Ef(h(\xi_n)) \rightarrow Ef(h(\xi)) \Rightarrow h(\xi_n) \rightarrow^d h(\xi)$

□

Лемма 1 <Слущкого>. Пусть $\xi_n \rightarrow^d \xi, \eta_n \rightarrow^d c = const$ - СВ. Тогда: $\xi_n + \eta_n \rightarrow^d \xi + c ; \xi_n \eta_n \rightarrow^d c\xi$

2. Вероятностно-статистическая модель

Определение 1. Пусть X - некот. наблюдение. Мн-во всех значений X наз. выборочным пространством и обозначается X .

X - результат случ. выбора элемента X с неизвестным распределением P .

Определение 2. Тройка (X, B_x, P) , где X - выборочное пространство, B_x - σ -алгебра на X , $P = \{P : P$ - вер.мера на $(X, B_x)\}$ - семейство распределений на (X, B_x) .

Если P параметризовано, т.е. $P = \{P_\theta : \theta \in O\}$, то модель (X, B_x, P) наз. параметрической. Обычно $(X, B_x) = (R^n, B(R^n))$.

Пример 2. Выборка. Пусть прибор работает некоторое случайное время. Все приборы однородны, а потому можно считать, что их времена работы - это независимые однородно распределенные СВ $\xi_1 \dots \xi_n \dots$. Пусть распр. времени работы определяется средним значением: $\Theta = E\xi_i$. Задача - оценить Θ ?

Пример 3. Линейная регрессия. Объект движется из положения Θ_1 в положение Θ_2 с нек. скоростью. Засаекаем его положение в нек. моменты времени. Известны результаты измерений положения объекта в моменты времени $t_1 \dots t_n$.

$$X_i = \Theta_1 + t_i \Theta_2 + \varepsilon_i$$

ε_i - ошибка измерения. Задача - оценить Θ_1, Θ_2

Определение 3. Набор $X_1 \dots X_n$ - независимых одинаково распределенных СВ с распр. P называется выборкой размера n из распределения P .

Определение 4. Пусть $X_1 \dots X_n$ - выборка. Тогда $\forall B \in B(R)$ обозначим $P_n^*(B) = U_n(B)/n$ где $U_n(B)$ - число элементов выборки, попавших в B , т.е. $P_n^* = \frac{\sum_{i=1}^n I\{X_i \in B\}}{n}$

Утверждение 1. Пусть $X_1 \dots X_n \dots$ - выборка неогр. размера из P_x . Тогда $\forall B \in B(R) : P_n^* \xrightarrow{a.s.} P_x(B)$

Доказательство. $P_n^*(B) = \frac{1}{n} \sum i = 1/n I\{X_i \in B\}$. Согласно УЗБЧ $P_n^* \xrightarrow{a.s.} EP_n^*(B) = EI\{X_i \in B\} = P(X_i \in B) = P_x(B)$ □

Определение 5. Пусть $X_1 \dots X_n$ - выборка. Тогда $F_n(x) = P_n^*(-\infty; x] = \frac{1}{n} \sum i = 1/n I\{X_i \leq x\}$

Теорема 9 <Гливленко-Кантелли>. Пусть $X_1 \dots X_n \dots$ - выборка неогр. размера с ф.р. $F(x)$, задана на вер. пр-ве (Ω, F, P) . Тогда $\sup_{x \in R} |F_n(x) - F(x)| \xrightarrow{a.s.} 0$

Доказательство. Пусть Q - мн-вол рац. чисел на R . $\forall \omega \in \Omega \longrightarrow |F_n(x, \omega) - F(x)|$ - непрерывн. справа, поэтому $\sup_{x \in R} |F_n(x, \omega) - F(x)| \geq \sup_{x \in Q} |F_n(x, \omega) - F(x)|$

Тогда $D_n(x) = \sup_{x \in Q} |F_n(x, \omega) - F(x)|$ - есть \sup счетного мн-ва случ. величин $D_n(\omega)$ - тоже СВ.

Пусть $N \in N$ Посмотрим $\forall k = 1 \dots N - 1 \longrightarrow X_{k,N} = \min\{x : F(x) \geq k/n\}$ Для удобства положим $x_{0,N} = -\infty, x_{N,N} = +\infty$ Пусть $x \in [x_{k,N}, x_{k+1,N}]$ $F_n(x) - F(x) \leq F_n(x_{k+1,N} - 0) - F(x_{k,N}) = (F_n(x_{k+1,N} - 0) - F(x_{k+1,N} - 0)) + (F(x_{k+1,N} - 0) - F(x_{k,N}))$ Аналогично, \dots $F_n(x) - F(x) \geq F_n(x_{k,N}) - F(x_{k,N}) + F(x_{k,N}) - F(x_{k+1,N} - 0) \geq F_n(x_{k,N}) - F(x_{k,N}) - 1/N$ Т.о. $\forall x \in R \longrightarrow |F_n(x) - F(x)| \leq \max_{0 \leq k \leq N-1, 1 \leq l \leq N} \{F_n(x_{k+1,N} - 0) - F(x_{k+1,N} - 0)\}, |F_n(x_{k,N})$