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

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

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

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

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

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

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

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 $\xi_n \to^p \xi$, то существует такая $\{\xi_n k\}$, что $\xi_n k \to^{a.s.} \xi$*

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

Теорема 6 <УЗБЧ>.
Пусть $\{\xi_n\}_{n=1}^{\infty}$ - нез. о.р. СВ с ограниченной дисперсией. Обозначим S_n (аналогично). Тогда $\frac{S_n - ES_n}{n} \to^{a.s.} 0$

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 $\frac{S_n - ES_n}{\sqrt{DS_n}} \to^d N(0, 1)$

Теорема 8 <о наследовании сх-ти>.
Пусть $\xi, \{\xi_n\}_{n=1}^{\infty}$ - случайные векторы размерности m.

1. *Если $\xi_n \to^{a.s.} \xi$ и $h(x) : R^m \to R^l$ 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 $h(\xi_n) \to^{a.s.} h(\xi)$*
2. *Если $\xi_n \to^p \xi$ и $h(x) : R^m \to R^l$ такова, что h непрерывна почти всюду относительно распределения ξ. Тогда $h(\xi_n) \to^p h(\xi)$*
3. *Если $\xi_n \to^d \xi$ и $h(x) : R^m \to R^l$ - непрерывна. Тогда $h(\xi_n) \to^d h(\xi)$*

Доказательство.

1. $1 \geq P(\lim_{n\to\infty} h(\xi_n) = h(\xi)) = P(\lim_{n\to\infty} h(\xi_n) = h(\xi), \xi \in B) \geq P(\lim_{n\to\infty} \xi_n = \xi, \xi \in B) = 1 \Rightarrow P(\lim_{n\to\infty} h(\xi_n) = h(\xi)) = 1$

2. Пусть $h(\xi_n) \not\to^p h(\xi)$. Тогда $\exists \varepsilon_0 > 0 : \exists$ подпослед. $\xi_n k, \exists \delta_0 > 0 : P(|h(\xi_n k) - h(\xi)| > \varepsilon_0) \geq \delta_0 \forall k$. Но $\xi_n k \to^p \xi \Rightarrow$ есть еще подпослед. $\xi_{nk_s} : \xi_{nk_s} \to^{a.s.} \xi, s \to \infty$

3. Возьмем $f(x) : R^m \to R^m$ -огр, непрер. ф-я. $Ef(h(\xi_n)) \to^? Ef(h(\xi))$. Но $f(h(x))$ - непрерывная ограниченная в R^n и $\xi_n \to^d \xi$. Отсюда $Ef(h(\xi_n)) \to Ef(h(\xi)) \Rightarrow h(\xi_n) \to^d h(\xi)$

□

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

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 неизвестгным 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. Засекаем его положение в нек. моменты времени. Известны результаты измерений положения объекта в моменты времени $t_1 \dots t_n$.

$$X_i = \Theta_1 + t_1\Theta_2 + \varepsilon_i$$

ε_i - ошибка измерения. Задача - оценить Θ_1, Θ_2

Определение 3. Набор $X_1 \dots X_n$ - независимых одинаково распределнных 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. $P_n^*(B) = \frac{1}{n} \sum i = 1nI\{X_i \in B\}$. 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 $F_n(x) = P_n^*(-\infty; x] = \frac{1}{n} \sum i = 1nI\{X_i \leq x\}$

Теорема 9 <Гливенко-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* $\sup_{x\in R} |F_n(x) - F(x)| \xrightarrow{a.s} 0$