Стационарные выборочные функции распределения, критерий Колмогорова (без вывода)

Луничкин Е.В.

Московский Физико-Технический Институт

21 мая 2018

О рядах

- Пусть на подмножестве $T \subseteq \mathbb{R}$, где $\mathbb{R}$ - множество действительных чисел, задана функция $x(t), t \in T$.
- Временной ряд: любое **счетное** множество X упорядоченных по времени значений величины x.

Выборочные функции распределения и моменты

- Множество принимаемых значений $x(t)$ ограничено, так как промежуток наблюдения ограничен
- Считаем ряд $x(t)$ равномерно ограниченным по времени
- Б.о.о. $D(x) = [0; 1]$

Выборочные функции распределения и моменты

Относительная частота события

Пусть в выборке объема T значение x_i случайной величины ξ встретилось $n_i(T)$ раз. Относительной частотой этого события называется отношение:

$$\nu_i(T) = \frac{n_i(T)}{T}. \tag{1}$$

Эмпирическое распределение

Совокупность величин $\nu_i(T)$ называется *эмпирическим распределением* или *выборочной плотностью функции распределения (ВПФР)* $f_T(x)$.

Эмпирическая функция распределения

Эмпирическая функция распределения

Эмпирической или *выборочной функцией распределения* называется ступенчатая неубывающая функция $F_T(x)$, определяемая по эмпирическому распределению частот по формуле:

$$F(x) \equiv F(k) = \sum_{i=1}^{k} \nu_i(T). \tag{2}$$

Эмпирическая функция распределения

Теорема Гливенко-Кантелли

ВФР $F_T(x)$ стационарной случайной величины равномерно по x сходится по вероятности к соответствующему распределению $F(x)$ генеральной совокупности при $T \to \infty$, т.е.

$$P\{\lim_{T\to\infty} \sup_x |F_T(x) - F(x)| = 0\} = 1. \tag{3}$$

Это означает, что

- эмпирические вероятности (1) сходятся к теоретическим;
- относительная частота $\nu_i(T)$ - несмещенная состоятельная оценка вероятности p_i.

Свойства оценок

Несмещенная оценка

Оценка $\tilde{\theta}_T$ величины θ называется *несмещенной*, если её матожидание равно оцениваемой величине:

$$\mathbb{E}[\tilde{\theta}_T] = \theta. \tag{4}$$

Состоятельная оценка

Оценка $\tilde{\theta}_T$ величины θ называется *состоятельной*, если она сходится по вероятности к оцениваемой величине:

$$P\{\lim_{T\to\infty} \tilde{\theta}_T - \theta = 0\} = 1, \tag{5}$$

т.е. $\forall \varepsilon > 0 P(|\tilde{\theta}_T - \theta| \leq \varepsilon) \to 1,\, T \to \infty$.

Эмпирическая функция распределения

Утверждение

Если ряд $x(t)$ стационарный, то при $T \to \infty$ разность $F_T(x) - F(x)$ распределена асимптотически нормально с параметрами $\mu = 0$, $\sigma^2 = \frac{F(x)(1-F(x))}{T}$,

т.е. при больших объемах выборки плотность вероятности величины $z = F_T(x) - F(x)$ является гауссовой и определяется выражением

$$P(z \in (a, a+\delta)) = \int_a^{a+\delta} f_G(z; \mu, \sigma) dz, f_G(z; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(z-\mu)^2}{2\sigma^2}) \tag{6}$$

Критерий Колмогорова

Из (6) следует, что величина уклонения $F_T(x) - F(x)$ неравномерна по x. Для получения равномерных оценок уклонения следует рассматривать статистику Колмогорова: $D_T = \sup_x |F_T(x) - F(x)|$. Распределение этой статистики дается теоремой Колмогорова.

Теорема Колмогорова

Если теоретическое распределение $F(x)$ генеральной совокупности непрерывно, то ВФР статистики $\sqrt{T}D_T$ сходится при $T \to \infty$ к функции Колмогорова $K(z)$:

$$\lim_{T\to\infty} P\{0 < \sqrt{T}\sup_x |F_T(x) - F(x)| < z\} = K(z) = \\ = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2z^2) \quad (7)$$

Статистика Смирнова

Обычно распределение генеральной совокупности не бывает известно. Тогда, в предположении, что оно существует, для изучения вопроса принадлежности двух ВФР одному и тому же распределению генеральной совокупности применяется статистика Смирнова $D_{m,n} = \sup_x |F_{1,m}(x) - F_{2,n}(x)|$, построенная по двум выборкам объемов m и n.

Относительно этой статистики справедливо следующее утверждение.

Утверждение

Пусть проводятся две независимых серии испытаний по составлению выборок объемов m и n из некоторой генеральной совокупности. Тогда ВФР статистики $\sqrt{\frac{mn}{m+n}} D_{m,n}$ сходится при $n \to \infty$ к функции Колмогорова в смысле формулы (7).