Этакое большое ничего и матстат

Белкин Дмитрий, U-1152
Бертыш Вадим, СПБГЭТУ «ЛЭТИ» 4373

15 июня 2016

Основные определения

Определение 1 (Статистический эксперимент)**.** *Тройка* $(\mathfrak{X}, \mathfrak{F}, \mathcal{P})$ *называется статистическим экспериментом*

- $\mathfrak{X}$ - *Множество результатов эксперимента*
- $\mathfrak{F}$ - *Совокупность наблюдаемых событий*
- $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ - *Семейство вероятностных распределений*

Дальше положим $\mathfrak{X} = \mathbb{R}^n$, $\mathfrak{F} = \sigma(\mathfrak{F}_1 \times \cdots \times \mathfrak{F}_n) = \mathfrak{B}_n$

Определение 2 (Статистика)**.** *Измеримая функция* $T : \mathfrak{X} \to E$ *называется статистикой*

Определение 3 (Подчиненная статистика)**.** *Статистика* T *называется подчиненной, если её распределение не зависит от параметра*

$$P_\theta(T \in A) = P_T(A)$$

Определение 4 (Достаточная статистика)**.** *Статистика* T *назвается достаточной, если условное распределение* X *при условии* T *не зависит от параметра*

$$P_\theta(X \in A|T) = P_{X|T}(A), \forall\theta \in \Theta$$

Подчиненная не содержит информации о параметре, достаточная содержит всю информацию о параметре

Определение 5 (Минимальная достаточная статистика)**.** *Достаточная статистика* T *называется минимальной, если,* $\forall T_1$ *достаточной* $\exists g : T = g(T_1)$

Использование МДС максимально редуцирует имеющиеся данные

Основные типы задач статистики

- Точечное оценивание (статистики $\delta : \mathfrak{X} \to \Theta$)
- Доверительное оценивание с уровнем доверия $1 - \alpha$ ($\mathcal{Y}$ - семейство подмножеств Θ)

  $$\Delta : \mathfrak{X} \to \mathcal{Y}$$

  такие, что $P_\theta(\theta \in \Delta(\vec{X})) \geq 1 - \alpha, \forall\theta \in \Theta$

- Проверка гипотез (принятие решений)
  $H : \theta \in \Theta_*, \Theta_* \subset \Theta$ - Гипотеза. Выдвигают $H_0 : \theta \in \Theta_0$ и $H_A : \theta \in \Theta$
  Решающее правило - критерий

  $$\phi : \mathfrak{X} \to [0; 1]$$

  $\phi(\vec{X})$ - вероятность выбрать альтернативу (отвергнуть H_0)

Асимптотический подход Пусть $(\mathfrak{X}^{(n)}, \mathfrak{F}^{(n)}, \mathcal{P}^{(n)})$ последовательность статистических экспериментов $\mathcal{P}^{(n)} = \{p_\theta^{(n)}, \theta \in \Theta\}$

Определение 6 (Состоятельность оценки)**.** *Точечная оценка $\delta^{(n)}(\vec{X})$ называется состоятельной, если*

$$\delta^{(n)}(\vec{X}) \xrightarrow{p_\theta} \theta, \forall \theta \in \Theta$$

Определение 7 (Сильная состоятельность оценки)**.** *Точечная оценка $\delta^{(n)}(\vec{X})$ называется сильно состоятельной, если*

$$\delta^{(n)}(\vec{X}) \xrightarrow[n \to \infty]{p_\theta = 1} \theta, \forall \theta \in \Theta$$

Определение 8 (Асимптотическая нормальность)**.** *Точечная оценка $\delta^{(n)}(\vec{X})$ называется асимптотически нормальной, если*

$$\sqrt{n}(\delta^{(n)}(\vec{X}) - \theta) \underset{P_\theta}{\Rightarrow} \mathcal{N}(0, \sigma^2(\theta))$$

Индикатор $\mathbb{1}$

Определение 9 (Точечная оценка)**.** *Статистика* $\delta(\vec{X})$, $\delta : \mathfrak{X} \to \Theta$ *называется точечной оценкой*

Определение 10 (Функция потерь)**.** *пусть* θ *реально значение параметра, тогда* $W(\delta(\vec{X}), \theta)$ *функция потерь, если*

- $W(\delta(\vec{X}), \theta) > 0, \forall \vec{X} \in \mathfrak{X}$
- $W(\theta, \theta) = 0$

Используют различные функции потерь (в дальнейшем используем функцию Гаусса)

$$W(\delta, \theta) = |\delta - \theta| \qquad \text{(Лаплас)}$$

$$W(\delta, \theta) = (\delta - \theta)^2 \qquad \text{(Гаусс)}$$

Определение 11 (Риск)**.** *Риском называют* $R(\delta, \theta) = \mathrm{E}$