Лекция 6.

Проверка статистических гипотез

Пусть $\vec{X} = (X_1, \ldots, X_n)$ из некоторого распределения F

Def. Гипотезой H называется предположение о распределении наблюдаемой случайной величины.

Доказать какое-то утверждение с помощью методов матстатистики невозможно - можно лишь с какой-то долей уверенности утверждать

Def. Гипотеза называется простой, если она однозначно определяет распределение: $H : F = F_1$, где F_1 - распределение известного типа с известными параметрами

В противном случае гипотеза называется сложной - она является объединением конечного или бесконечного числа простых гипотез

Например, «величина X принадлежит нормальному распределению» - сложная гипотеза, а «величина X принадлежит нормальному распределению с матожиданием $a = 1$ и дисперсией $\sigma^2 = 1$» - простая

В общем случае работаем со схемой из двух или более гипотез. В ходе проверки принимается ровна одна из них. Мы ограничимся самой простой схемой из 2 гипотез: H_0 - основная (нулевая) гипотеза, $H_1 = \overline{H_0}$ - альтернативная (конкурирующая) гипотеза, состоящая в том, что основная гипотеза неверна

Основная гипотеза H_0 принимается или отклоняется при помощи статистики критерия K

$K(X_1, \ldots, X_n) \longrightarrow \mathbb{R} = \overline{S} \cup S \longrightarrow (H_0, H_1)$

$$\begin{cases} H_0, & \text{если } K(X_1, \ldots, X_n) \in \overline{S} \\ H_1, & \text{если } K(X_1, \ldots, X_n) \in S \end{cases}$$

Вместо «гипотеза доказана» лучше употреблять «гипотеза принимается/отвергается»

Область S называется критической областью, а точка $t_{\text{кр}}$ на границе областей называется критической

Def. Ошибка первого рода состоит в том, что H_0 отклоняется, хотя она верна. Аналогично, ошибка второго рода состоит в том, что H_1 отклоняется, хотя она верна.

Def. Вероятность α ошибки первого рода называется уровнем значимости критерия. Вероятность ошибки второго рода обозначаем β. Мощностью критерия называется вероятность $1 - \beta$ (вероятность недопущения ошибки второго рода)

Ясно, что критерий будет тем лучше, чем меньше вероятности ошибок α и β. При увеличении объема выборки уменьшаются обе вероятности. При фиксированном объему попытки уменьшить одну вероятность увеличат другую

Одним из способов является фиксация одной вероятности (принято α) и уменьшение другой

Построение критериев согласия

Def. Говорят, что критерий K является критерием асимптотического уровня ε, если вероятность ошибки первого рода $\alpha \underset{n\to\infty}{\longrightarrow} \varepsilon$

Def. Критерий K для проверки гипотезы H_0 называется состоятельным, если вероятность ошибки второго рода $\beta \underset{n\to\infty}{\longrightarrow} 0$

Def. Критерием согласия уровня ε называем состоятельный критерий асимптотического уровня ε

Обычно критерий согласия строится по следующей схеме: берется статистика $K(X_1, \dots, K_n)$, обладающая свойствами:

1. Если H_0 верна, то $K(X_1, \dots, X_n) \rightrightarrows Z$, где Z - известное распределение
2. Если H_0 неверна, то есть верна H_1, то $K(X_1, \dots, X_n) \xrightarrow[n\to\infty]{p} \infty$ (достаточно сильно отклоняться от распределения Z)

Построенный таким образом критерий является критерием согласия, то есть обладает свойствами

1. критерия асимптотического уровня
2. состоятельного критерия

Пусть $t_{\text{кр}}$ - критическая точка такая, что $P(|Z| \geq t_{\text{кр}}) = \varepsilon$ - заданный уровень ошибки первого рода

$$\begin{cases} H_0, & \text{если } |K| < t_{\text{кр}} \\ H_1, & \text{если } |K| \geq t_{\text{кр}} \end{cases}$$

1. Тогда $\alpha = P(|K| \geq t_{\text{кр}} \mid H_0) = 1 - P(|K| < t_{\text{кр}} \mid H_0) = 1 - (F_K(t_{\text{кр}}) - F_K(-t_{\text{кр}})) \xrightarrow[n\to\infty]{\substack{\text{т.к. при верной } H_0 \\ F_K(x) \underset{n\to\infty}{\longrightarrow} F_Z(x)}} 1 - (F_Z(t_{\text{кр}}) - F_Z(-t_{\text{кр}})) = P(|Z| \geq t_{\text{кр}}) = \varepsilon$
2. Если H_1 верна, то $|K| \xrightarrow{p} \infty$, то есть $\forall C$ $P(|K| > C \mid H_1) \xrightarrow{p} 1 \implies \beta = P(|K| < C \mid H_1) \xrightarrow{p} 0$

Гипотеза о среднем нормальной совокупности при известной дисперсии

Пусть $\vec{X} = (X_1, \dots, X_n)$ из $N(a, \sigma^2)$, причем σ^2 известен.

Проверяется гипотеза, что $H_0 : a = a_0$, против $H_1 : a \neq a_0$ для уровня значимости α

1. По пункту 1 теоремы, если $H_0 : a = a_0$ верна, то $K = \sqrt{n}\dfrac{\overline{x} - a_0}{\sigma} = \sqrt{n}\dfrac{\overline{x} - a}{\sigma} \in N(0, 1)$

2. Если верна $H_1 : a \neq a_0$, то $|K| = \sqrt{n}\left|\dfrac{\overline{x} - a_0}{\sigma}\right| = \sqrt{n}\left|\dfrac{\overline{x} - a}{\sigma} + \dfrac{a - a_0}{\sigma}\right| =$

$$= \left| \underbrace{\sqrt{n}\frac{\overline{x} - a}{\sigma}}_{\substack{\in N(0,1),\text{ограничен} \\ \text{по вероятности}}} + \underbrace{\sqrt{n}}_{\to\infty} \underbrace{\frac{a - a_0}{\sigma}}_{\text{const}} \right| \xrightarrow{p} \infty$$

Для уровня значимости α находим $t_{\text{кр}}$ такую, что $\alpha = P(|K| \geq t_{\text{кр}} \mid H_0) = P(|Z| \geq t_{\text{кр}}) \Longrightarrow P(|Z| < t_{\text{кр}}) = 2F_0(t_{\text{кр}}) - 1 = 1 - \alpha$

$F_0(t_{\text{кр}}) = 1 - \dfrac{\alpha}{2}$ - то есть $t_{\text{кр}}$ - квантиль стандартного нормального распределения уровня $1 - \dfrac{\alpha}{2}$

$$\begin{cases} H_0, & \text{если } |K| < t_{\text{кр}} \\ H_1, & \text{если } |K| \geq t_{\text{кр}} \end{cases}$$

Гипотеза о среднем нормальной совокупности при неизвестной дисперсии

1. По пункту 4 основной теоремы, если $H_0 : a = a_0$ верна, то $K = \sqrt{n}\dfrac{\overline{x} - a_0}{S} = \sqrt{n}\dfrac{\overline{x} - a}{S} \in T_{n-1}$
2. Если верна $H_1 : a \neq a_0$, то $|K| = \sqrt{n}\left|\dfrac{\overline{x} - a_0}{S}\right| = \sqrt{n}\left|\dfrac{\overline{x} - a}{S} + \dfrac{a - a_0}{S}\right| =$

$$= \left| \underbrace{\sqrt{n}\frac{\overline{x} - a}{S}}_{\substack{\in T_{n-1},\text{ограничен} \\ \text{по вероятности}}} + \underbrace{\sqrt{n}}_{\to\infty} \underbrace{\frac{a - a_0}{S}}_{\text{const}} \right| \xrightarrow{p} \infty$$

Аналогично получаем $t_{\text{кр}}$ - квантиль распределения T_{n-1} уровня $1 - \dfrac{\alpha}{2}$

Доверительные интервалы как критерии гипотез по параметрам распределения

Пусть $(X_1, \ldots, X_n)$ из F_θ, где F_θ - распределение известного типа с неизвестным параметром θ

Проверяется гипотеза, что $H_0 : \theta = \theta_0$, против $H_1 : \theta \neq \theta_0$

Допустим, что для θ построен доверительный интервал $(\theta_\gamma^-, \theta_\gamma^+)$, то есть $P(\theta_\gamma^- < \theta < \theta_\gamma^+) = \gamma$.

Тогда критерий $\begin{cases} H_0, & \text{если } \theta_0 \in (\theta_\gamma^-, \theta_\gamma^+) \\ H_1, & \text{если } \theta_0 \notin (\theta_\gamma^-, \theta_\gamma^+) \end{cases}$ будет уровня $\alpha = 1 - \gamma$

$\alpha = P(\theta_0 \notin (\theta_\gamma^-, \theta_\gamma^+) \mid H_0) = 1 - P(\theta_0 \in (\theta_\gamma^-, \theta_\gamma^+) \mid X \in F_{\theta_0}) = 1 - \gamma$

Поэтому доверительные интервалы можно использовать для проверки гипотез

Но почему в схеме $\begin{cases} H_0 : a = \overline{x} \\ H_1 : a \neq \overline{x} \end{cases}$ основная гипотеза всегда верна, тогда как выборочно среднее на практике почти всегда не равняется матожиданию. Потому что ...

> А вот нефиг такие гипотезы вообще выдвигать
>
> © Блаженов А. В.

Критерий вероятности появления события

$\exists P(A) = p$ - вероятность успеха при одном испытании. При достаточно большом количестве испытаний n событие A появилось m раз. Проверяется $H_0 : p = p_0$ против $H_1 : p \neq p_0$

В качестве статистики критерия возьмем величину $K = \dfrac{m - np_0}{\sqrt{np_0q_0}}$

1. Если H_0 верна, то $K = \dfrac{m - np}{\sqrt{npq}} \rightrightarrows N(0,1)$ по ЦПТ
2. Lab.

Из тех же соображений $t_{\text{кр}}$ - квантиль $N(0,1)$ уровня $1 - \dfrac{\alpha}{2}$

$$\begin{cases} H_0 : p = p_0, & \text{если } |K| < t_{\text{кр}} \\ H_1 : p \neq p_0 & \text{если } |K| \geq t_{\text{кр}} \end{cases}$$

Ex. При посеве 4000 семян 970 всходов оказались рецессивного цвета, а 3030 - доминантного. Проверим гипотезу $H_0 : p = \dfrac{1}{4}$ - Мендель прав, против $H_1 : p \neq \dfrac{1}{4}$ - Мендель не прав, для уровня значимости - 0.05

$$K = \frac{m - np_0}{\sqrt{np_0q_0}} = \frac{970 - 4000 \cdot \frac{1}{4}}{\sqrt{4000\frac{1}{4}\frac{3}{4}}} \approx -1.095$$

Так как $|K| = 1.095 < 1.96 = t_{\text{кр}}$, то $H_0 : p = \dfrac{1}{4}$ верна