GeekBrains

Теория вероятностей и математическая статистика

Вебинары

GeekBrains

Урок 5

Теория вероятностей и математическая статистика

Проверка статистических гипотез. P-значения. Доверительные интервалы. A/B тестирование

На этом уроке мы изучим:

1. Что такое статистическая гипотеза.
2. Нулевые и альтернативные гипотезы.
3. Статистические критерии для проверки гипотез.
4. P-value
5. Доверительные интервалы.

Статистическая гипотеза

Предположение о виде распределения и свойствах случайной величины, которое можно подтвердить или опровергнуть на основании имеющихся данных

Примеры:

- Математическое ожидание случайной величины равно 10
- Случайная величина имеет нормальное распределение
- Две случайные величины имеют одинаковое математическое ожидание.

Этапы проверки гипотез

1. Формулируются нулевая и альтернативная гипотезы.

Статистическая гипотеза

Нулевая гипотеза – утверждение о свойствах генеральной совокупности, которое кажется правдоподобным, но требует проверки, та гипотеза, которую проверяют (H_0)

Альтернативная гипотеза – гипотеза, противоречащая нулевой (H_1)

<u>Пример:</u> Имеется станок, изготавливающий шарики для подшипников, который настроен делать шарики с диаметром 1 мм. На основании выборки из значений диаметров таких шариков мы можем проверить, правильно ли станок откалиброван

H_0 - математическое ожидание диаметра шарика равно 1 мм

H_1 – математическое ожидание диаметра шарика не равно 1 мм

Альтернативная гипотеза

$H_0\colon \mu = \mu_0$

1. Двухсторонние $H_1\colon \mu \neq \mu_0$
2. Левосторонние $H_1\colon \mu < \mu_0$
3. Правосторонние $H_1\colon \mu > \mu_0$

Альтернативная гипотеза – это дополнение к нулевой (хотя бы одна из них будет верна)

Этапы проверки гипотез

1. Формулируются нулевая и альтернативная гипотезы.
2. Задаётся некоторая статистика (функция от выборки) S(X), которая в условиях справедливости нулевой гипотезы H0 имеет известное распределение (в частности, известна её функция распределения FS(x) = P(S < x)).

Выбор статистики S

Для проверки гипотез относительно математического ожидания нормально распределенной случайной величины *с известной дисперсией* используется Z-статистика

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

В предположении верности нулевой гипотезы z-статистика имеет стандартное нормальное распределение (0, 1)

Если дисперсия не известна, используется t-статистика

$$t = \frac{\overline{X} - \mu}{\sigma_X / \sqrt{n}}$$

В предположении верности нулевой гипотезы t-статистика имеет распределение Стьюдента

GeekBrains

Выбор статистики S

MironkindY

Нулевая гипотеза H_0	Дополнительные условия	СТАТИСТИКА КРИТЕРИЯ (выборочная характеристика)	Используемые распределение и таблица (уч. МС)	Конкур. гип. H_1	Критическая область	формулы для нахождения её границ	Гипотеза H_0 **не отвергается**, если:
1. $H_0: \mu = \mu_0$	$X \in N(\mu;\sigma)$, σ^2 известна	$t_н = \frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}$	Нормальный закон Функция Лапласа $\Phi(t)$ таб.1	$\mu_1 > \mu_0$	ПКО	$\Phi(t_{кр}) = 1-2\alpha$	$t_н \le t_{кр}$
				$\mu_1 < \mu_0$	ЛКО		$t_н \ge -t_{кр}$
				Мощность критерия:		$1-\beta = \frac{1}{2}\left[1+\Phi\left(\frac{\lvert\mu_1-\mu_0\rvert}{\sigma}\sqrt{n}-t_{кр}\right)\right]$	
				$\mu_1 \ne \mu_0$	ДКО	$\Phi(t_{кр}) = 1-\alpha$	$\lvert t_н \rvert \le t_{кр}$
2. $H_0: \mu = \mu_0$	$X \in N(\mu;\sigma)$, σ^2 неизвестна	$t_н = \frac{\bar{x}-\mu_0}{S}\sqrt{n-1}$	Стьюдента $St(t,\nu)$ $\nu = n-1$ таб.2	$\mu_1 > \mu_0$	ПКО	$St(t_{кр},\nu) = 2\alpha$	$t_н \le t_{кр}$
				$\mu_1 < \mu_0$	ЛКО		$t_н \ge -t_{кр}$
				Мощность критерия:		$1-\beta = 1-\frac{1}{2}St\left(\frac{\lvert\mu_1-\mu_0\rvert}{S}\sqrt{n-1}-t_{кр}; n-1\right)$	
				$\mu_1 \ne \mu_0$	ДКО	$St(t_{кр},\nu) = \alpha$	$\lvert t_н \rvert \le t_{кр}$
3. $H_0: \mu_x = \mu_y$	$X \in N(\mu_x;\sigma_x)$, $Y \in N(\mu_y;\sigma_y)$, σ_x^2 и σ_y^2 известны	$t_н = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}}}$	Нормальный закон Функция Лапласа $\Phi(t)$ таб.1	$\mu_x > \mu_y$	ПКО	$\Phi(t_{кр}) = 1-2\alpha$	$t_н \le t_{кр}$
				$\mu_x < \mu_y$	ЛКО		$t_н \ge -t_{кр}$
				$\mu_x \ne \mu_y$	ДКО	$\Phi(t_{кр}) = 1-\alpha$	$\lvert t_н \rvert \le t_{кр}$
4. $H_0: \mu_x = \mu_y$	$X \in N(\mu_x;\sigma_x)$, $Y \in N(\mu_y;\sigma_y)$, σ_x^2 и σ_y^2 неизвестны, но равны (ПРОВЕРИТЬ!)	$t_н = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{n_x S_x^2+n_y S_y^2}{n_x+n_y-2}}} \cdot \sqrt{\frac{n_x n_y}{n_x+n_y}}$	Стьюдента $St(t,\nu)$ $\nu = n_x+n_y-2$ таб.2	$\mu_x > \mu_y$	ПКО	$St(t_{кр},\nu) = 2\alpha$	$t_н \le t_{кр}$
				$\mu_x < \mu_y$	ЛКО		$t_н \ge -t_{кр}$
				$\mu_x \ne \mu_y$	ДКО	$St(t_{кр},\nu) = \alpha$	$\lvert t_н \rvert \le t_{кр}$
5. $H_0: \sigma^2 = \sigma_0^2$	$X \in N(\mu;\sigma)$	$\chi_н^2 = \frac{nS^2}{\sigma_0^2}$	Пирсона χ^2 $\nu = n-1$ таб.3	$\sigma_1^2 > \sigma_0^2$	ПКО	$P(\chi^2 > \chi_{кр}^2(\alpha,\nu)) = \alpha$	$\chi_н^2 \le \chi_{кр}^2$
				Мощность критерия:		$1-\beta = P(\chi^2 > \frac{\sigma_0^2}{\sigma_1^2}\chi_{кр}^2(\alpha;n-1))$	
				$\sigma_1^2 < \sigma_0^2$	ЛКО	$P(\chi^2 > \chi_{кр}^2(1-\alpha,\nu)) = 1-\alpha$	$\chi_н^2 \ge \chi_{кр}^2$
				Мощность критерия:		$1-\beta = 1-P(\chi^2 > \frac{\sigma_0^2}{\sigma_1^2}\chi_{кр}^2(1-\alpha;n-1))$	
				$\sigma_1^2 \ne \sigma_0^2$	ДКО	$P(\chi_{кр.л.}^2) = 1-\frac{\alpha}{2}$; $P(\chi_{кр.п.}^2) = \frac{\alpha}{2}$	$\chi_{кр\,л}^2 \le \chi_н^2 \le \chi_{кр\,п}^2$

Выбор статистики S

H_0	Доп. условия	СТАТИСТИКА КРИТЕРИЯ	Распределение	H_1	Критическая область		H_0 не отвергается, если
6. $H_0: \sigma_1^2 = \sigma_2^2$	$X \in N(\mu_x;\sigma_x)$ $Y \in N(\mu_y;\sigma_y)$ $\hat{S}_1^2 > \hat{S}_2^2$ $S_1 = max(S_x;S_y)$	$F_н = \dfrac{\hat{S}_1^2}{\hat{S}_2^2}$ $\hat{S}_i^2 = \dfrac{n_i}{n_i - 1} S_i^2$	Фишера-Снедекора $F(\alpha, v_1, v_2)$ $v_i = n_i - 1$ таб.4	$\sigma_1^2 > \sigma_2^2$	ПКО	$P(F > F_{кр}(\alpha, \nu_1, \nu_2)) = \alpha$	$F_н \le F_{кр}$
7. $H_0: \sigma_1^2 = \sigma_2^2 = ..$ $... = \sigma_k^2$ k - число генеральных совокупностей	$X_i \in N(\mu_i;\sigma_i)$ $n_1 \neq n_2 \neq ...$ $... \neq n_k$	**Критерий Бартлетта** $\chi_н^2 = \dfrac{v_k \cdot \ln\hat{S}_{ср}^2 - \sum\limits_{i=1}^{k}\left(v_i \cdot \ln\hat{S}_i^2\right)}{1 + \dfrac{1}{3\cdot(k-1)}\left(\left(\sum\limits_{i=1}^{k}\dfrac{1}{v_i}\right) - \dfrac{1}{v_k}\right)}$ $v_i = n_i - 1;$ $v_k = \sum\limits_{i=1}^{k} v_i$ $\hat{S}_i^2 = \dfrac{n_i}{n_i - 1} S_i^2;$ $\hat{S}_{ср}^2 = \dfrac{1}{v_k}\sum\limits_{i=1}^{k}\left(\hat{S}_i^2 \cdot v_i\right)$	Пирсона χ^2 $\nu = k - 1$ таб.3		ПКО	$P(\chi^2 > \chi_{кр}^2(\alpha, \nu)) = \alpha$	$\chi_н^2 \le \chi_{кр}^2$
8. $H_0: \sigma_1^2 = \sigma_2^2 = ..$ $... = \sigma_k^2$ k - число ген.сов.	$X_i \in N(\mu_i;\sigma_i)$ $n_1 = n_2 = ...$ $... = n_k = n$	**Критерий Кохрана** $G_H = \dfrac{\max(\hat{S}_i^2)}{\sum\limits_{i=1}^{k}\hat{S}_i^2} = \dfrac{\max(S_i^2)}{\sum\limits_{i=1}^{k} S_i^2}$	G-распределение $G(\alpha, \nu, k)$ $\nu = n - 1$ таб.9		ПКО	$P(G > G_{кр}(\alpha, \nu, k)) = \alpha$	$G_н \le G_{кр}$
9. $H_0: p = p_0$	Биномиал. распред., $n \to \infty$ $(n > 30)$	$t_н = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 \cdot (1 - p_0)}{n}}}$ $\hat{p} = \dfrac{m}{n}$	Нормальный закон Функция Лапласа $\Phi(t)$ таб.1	$p > p_0$ $p < p_0$ $p \neq p_0$	ПКО ЛКО ДКО	$\Phi(t_{кр}) = 1 - 2\alpha$ $\Phi(t_{кр}) = 1 - \alpha$	$t_н \le t_{кр}$ $t_н \ge -t_{кр}$ $\lvert t_н \rvert \le t_{кр}$
10. $H_0: p_1 = p_2 = ...$ $... = p_k$ k - число ген.сов.	Биномиальное распред., $n \to \infty$ $(n > 30)$	$\chi_н^2 = \sum\limits_{i=1}^{k}\dfrac{\left(m_i - n_i \cdot \hat{p}\right)^2}{n_i \cdot \hat{p} \cdot (1 - \hat{p})} = \dfrac{\sum\limits_{i=1}^{k}\left(\hat{p}_i - \hat{p}\right)^2 \cdot n_i}{\hat{p} \cdot \left(1 - \hat{p}\right)}$ $\hat{p} = \sum\limits_{i=1}^{k} m_i \Big/ \sum\limits_{i=1}^{k} n_i$ $\quad \hat{p}_i = \dfrac{m_i}{n_i}$	Пирсона χ^2 $\nu = k - 1$ таб.3		ПКО	$P(\chi^2 > \chi_{кр}^2(\alpha, \nu)) = \alpha$	$\chi_н^2 \le \chi_{кр}^2$
11. $H_0: p_{1j} = p_{2j} = ...$ $... = p_{kj} = p_j$ $j = 1 \ldots h$ – группы	Полиномиальное распределение $N \to \infty$ $(N > 30)$	$\chi_н^2 = \sum\limits_{j=1}^{h}\sum\limits_{i=1}^{k}\dfrac{\left(m_{ij} - n_i \cdot \hat{p}_j\right)^2}{n_i \cdot \hat{p}_j}$ $\hat{p}_j = \sum\limits_{i=1}^{k} m_{ij} \Big/ N;$ $N = \sum\limits_{i=1}^{k} n_j = \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{h} m_{ij}$	Пирсона χ^2 $\nu = (k-1)\cdot(h-1)$ таб.3		ПКО	$P(\chi^2 > \chi_{кр}^2(\alpha, \nu)) = \alpha$	$\chi_н^2 \le \chi_{кр}^2$

Этапы проверки гипотез

1. Формулируются нулевая и альтернативная гипотезы.
2. Задаётся некоторая статистика (функция от выборки) $S(X)$, которая в условиях справедливости нулевой гипотезы $H0$ имеет известное распределение (в частности, известна её функция распределения $FS(x) = P(S < x)$).
3. Фиксируется уровень значимости α — допустимая для данной задачи вероятность ошибки первого рода (чаще всего 0.01, 0.05 или 0.1)

Ошибки первого и второго рода

Ошибка первого рода (false positive)- это отказ от нулевой гипотезы, несмотря на то, что она верна

Ошибка второго рода (false negative)- это принятие нулевой гипотезы, хотя она не верна

Между вероятностями ошибок первого и второго рода приходится балансировать , т.е. уменьшение одной вероятности приводит к увеличению другой

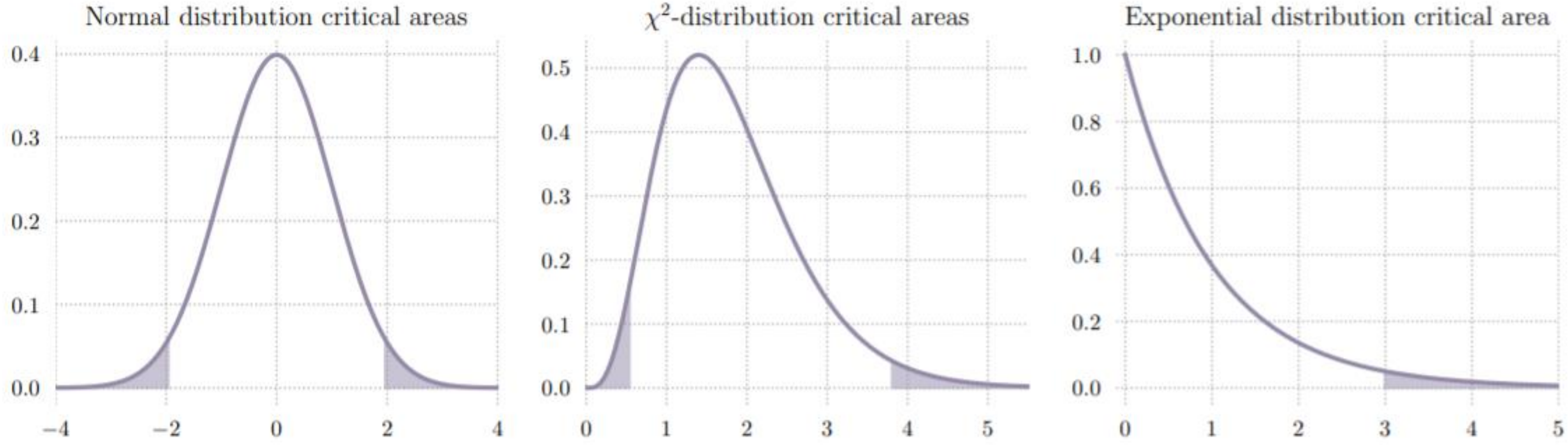
Уровень значимости α — вероятность отвергнуть верную нулевую гипотезу

Этапы проверки гипотез

1. Формулируются нулевая и альтернативная гипотезы.
2. Задаётся некоторая статистика (функция от выборки) $S(X)$, которая в условиях справедливости нулевой гипотезы $H0$ имеет известное распределение (в частности, известна её функция распределения $FS(x) = P(S < x)$).
3. Фиксируется уровень значимости α — допустимая для данной задачи вероятность ошибки первого рода (чаще всего 0.01, 0.05 или 0.1).
4. Определяется критическая область $\Omega\alpha$, такая, что $P(S \in \Omega\alpha|H0) = \alpha$.
5. Проводится статистический тест : для конкретной выборки X считается значение $S(X)$, и если оно принадлежит $\Omega\alpha$, то заключаем, что данные противоречат гипотезе $H0$, и принимается гипотеза $H1$.

Критическая область

$P(S \in \Omega_\alpha | H_0) = \alpha$



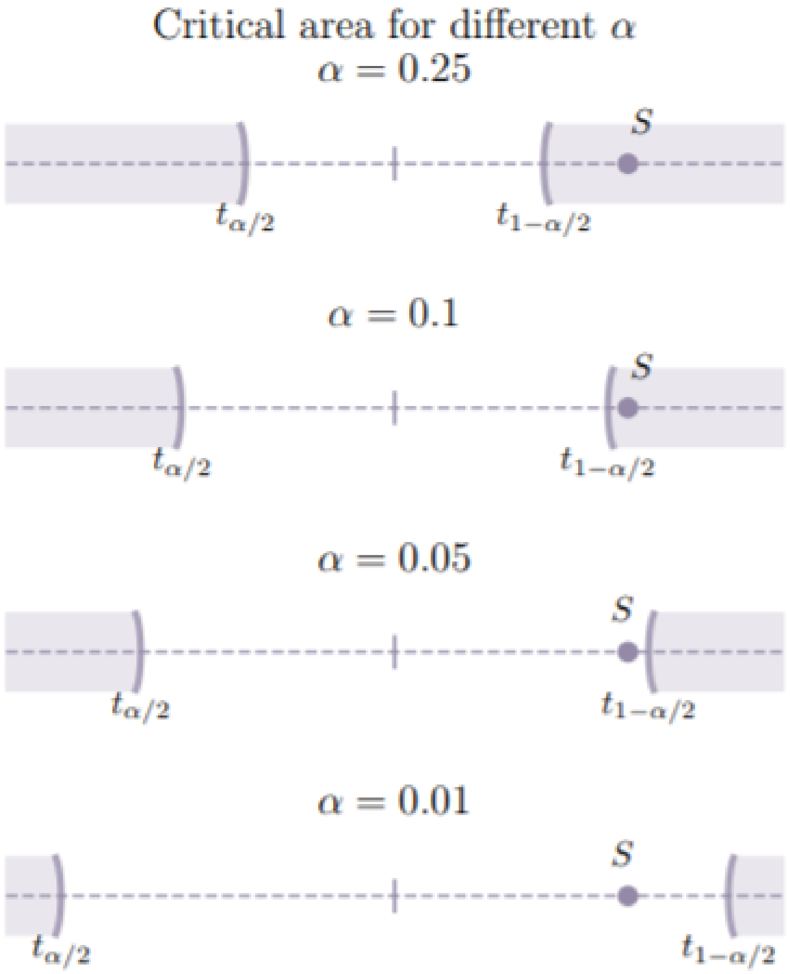
- ЛКО $\Omega_\alpha = (-\infty, t_\alpha)$
- ПКО $\Omega_\alpha = (t_{1-\alpha}, \infty)$
- ДКО $\Omega_\alpha = \left(-\infty, t_{\alpha/2}\right) \cup \left(t_{1-\alpha/2}, \infty\right)$

P-value

P-value — это такое значение α, при котором значение статистики попадает ровно на границу критической области

p-value > α ⇒ нулевая гипотеза не отвергается на уровне значимости α

p-value < α ⇒ нулевая гипотеза отвергается с вероятностью ошибки α



Доверительный интервал

Доверительный интервал — это интервал, который с некоторой вероятностью (заданной заранее) содержит значение оцениваемого параметра

Пусть задано число p, называемое уровнем доверия или доверительной вероятностью . Доверительным интервалом для параметра θ называется пара статистик L и U, таких, что

$$P(L \leq \theta \leq U) = p$$

Доверительный интервал

Пусть дана выборка X из нормально распределённой случайной величины с известной дисперсией. Построим доверительный интервал для математического ожидания μ с доверительной вероятностью p

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

$$P\left(t_{\alpha/2} \leq Z \leq t_{1-\alpha/2}\right) = p,$$

$$P\left(t_{\alpha/2} \leq \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \leq t_{1-\alpha/2}\right) = p$$

$$P\left(t_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \overline{X} - \mu \leq t_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) = p$$

$$P\left(\overline{X} + t_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) = p$$

Доверительный интервал

Если дисперсия не известна, используем t-статистику

$$t = \frac{\overline{X} - \mu}{\sigma_X / \sqrt{n}}$$

$$P\left(t_{\alpha/2,\, n-1} \leq t \leq t_{1-\alpha/2,\, n-1}\right) = p,$$

$$P\left(\overline{X} + t_{\alpha/2,\, n-1} \cdot \frac{\sigma_X}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{1-\alpha/2,\, n-1} \cdot \frac{\sigma_X}{\sqrt{n}}\right) = p$$

Итоги

1. Что такое статистическая гипотеза.
2. Нулевые и альтернативные гипотезы.
3. Статистические критерии для проверки гипотез.
4. P-value
5. Доверительные интервалы.