Проверка статистических гипотез 2

Н.П. Алексеева,

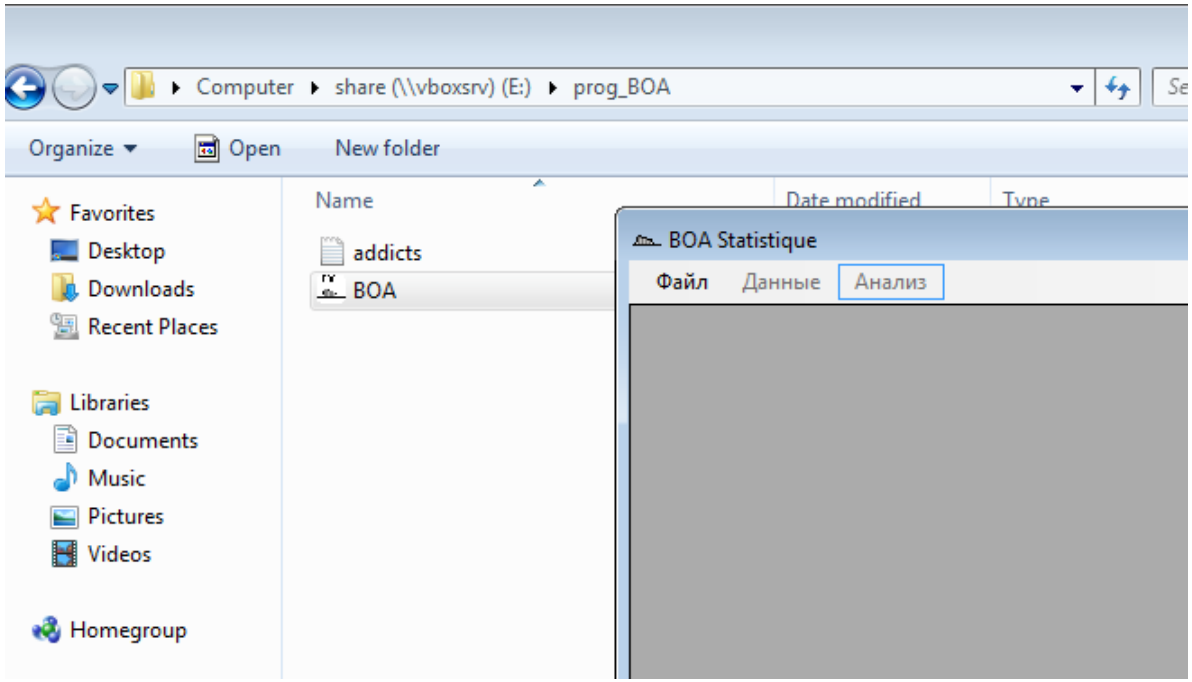
СПбГУ, математико-механический факультет

2020 г.

Программа BOA Statistique

http://mystatbook.narod.ru/minimumA.zip

По указанной ссылке получаем архивный файл, распаковываем доступными средствами. В результате должна быть установлена папка, в которой находится программа BOA и данные addicts.txt. Программа под Windows.



При загрузки программы появляется окно. Выбираем "Файл "Загрузить "addicts.txt".

Описание категориальных признаков

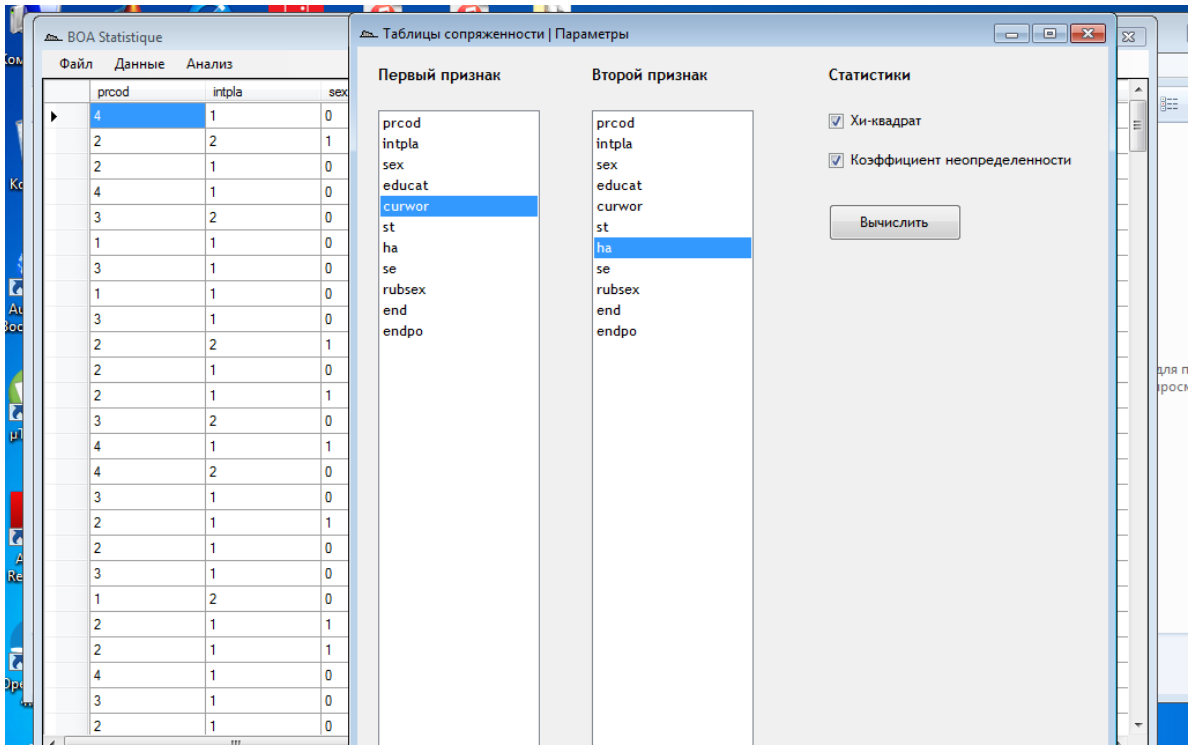
	Кодировка	расшифровка признака
1	prcod	лечебный препарат: 1 – ноутрексон и плацебо, 2 – плацебо и плацебо, 3 – ноутрексон и антидепрессант, 4 – плацебо и антидепрессант
2	intpla	место лечения: 1 – амбулаторно, 2 – в диспансере
3	sex	пол: 1 – женский, 0 – мужской
4	age	возраст
5	educat	образование: 1 – 8 классов, 2 – средняя школа, 3 – неполное высшее, 4 – высшее
6	curwor	занятость: 0 – не работает, 1 – работает
17	st	использование стимулирующих средств
18	ha	использование препаратов, вызывающих галлюцинации
19	se	использование успокоительных средств
26	end	итог лечения: 1 – программа лечения выполнена, 0 – программа сорвана

Пример задания

Вопросы по проверке независимости качественных признаков	да-нет	P	зн.-незн.
1. Не занятые на работе или учебе больные чаще используют препараты, вызывающие галлюцинации.	23%, 26% нет	0.5963	не значимо $\alpha = 0.05$
2. Среди женщин реже случаи срыва программы.	73%, 71% да	0.82	не значимо
3. Прием стимулирующих средств является препятствием для выполнения программы лечения.	29%, 15%	0.0778	значимо для $\alpha = 0.1$

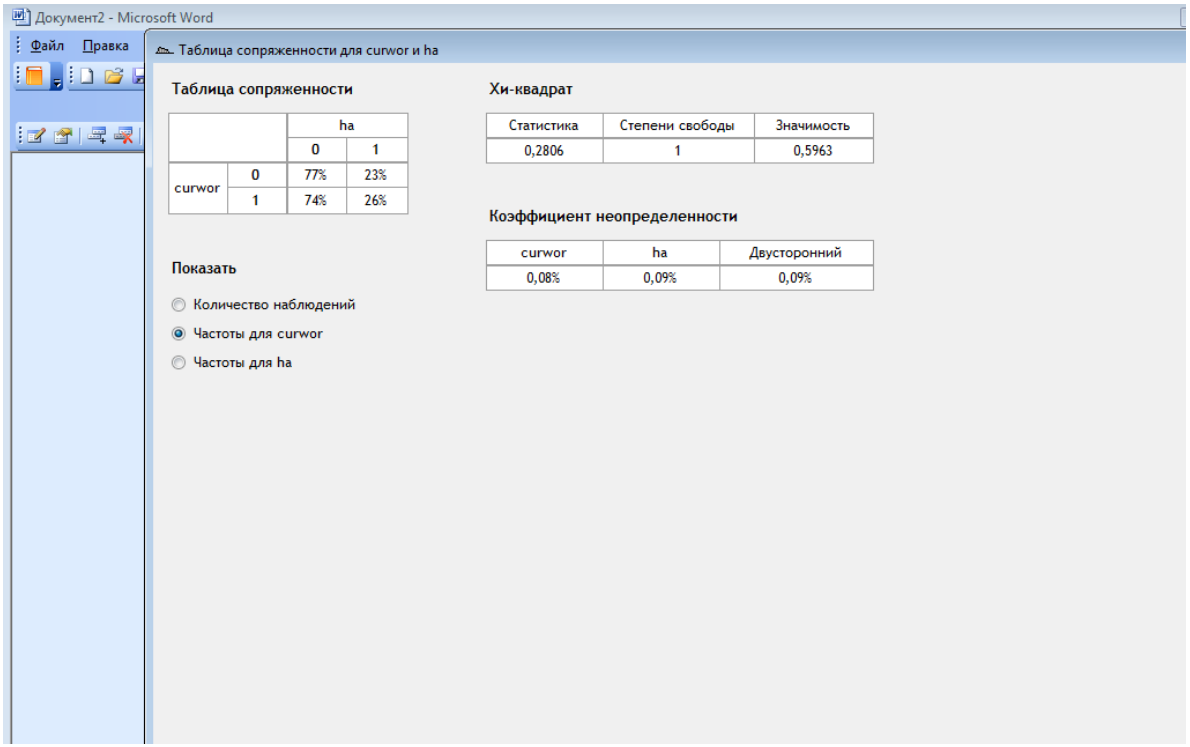
Пример решения задачи 1

Не занятые на работе или учебе больные чаще используют препараты, вызывающие галлюцинации.

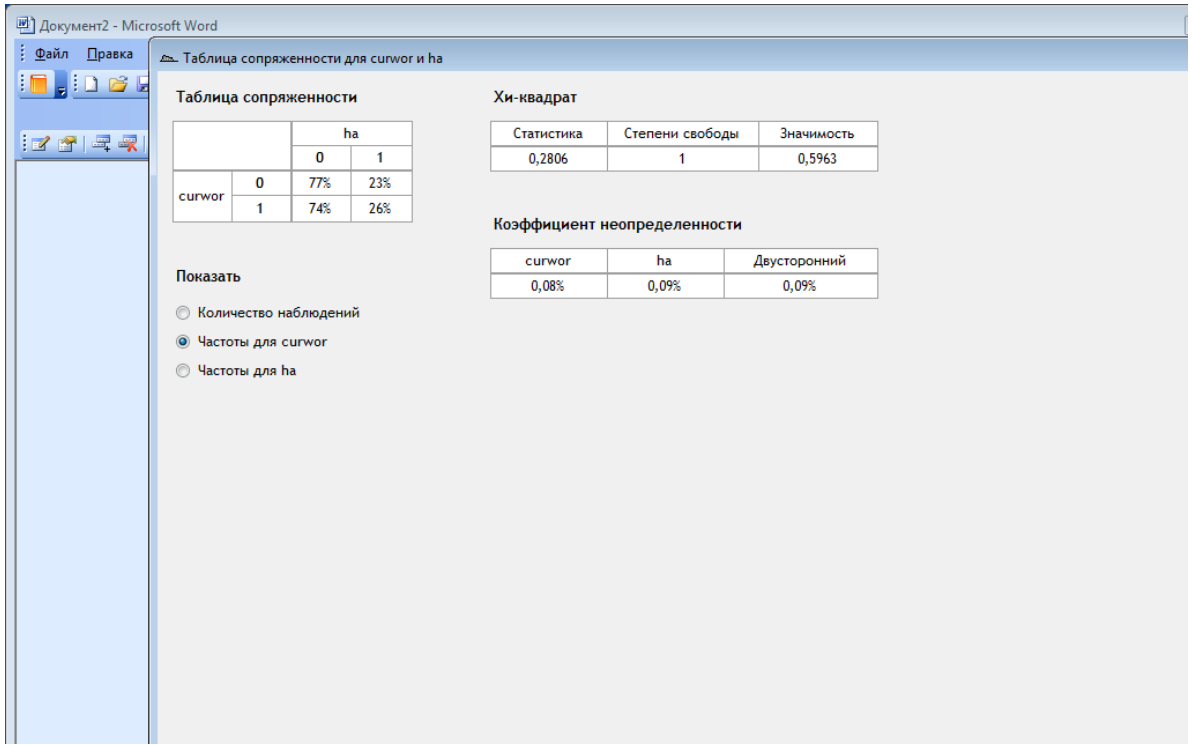


Анализ - Таблицы сопряженности. Выбираем из списков соответствующие переменные и кнопку "Вычислить".

Пример решения задачи 1



Пример решения задачи 1



Среди незанятых 23% употребляют галлюциногены, среди занятых 26%. Значимость различия p = 0.5963. Различие можно объяснить случайностью.

Задания на проверку

Вопросы по проверке независимости качественных признаков	да-нет	P	зн.-незн.
1. Прием успокоительных средств является препятствием для выполнения программы.			
2. Больные, принимающие вызывающие галлюцинации препараты, хуже выполнили программу.			
3. Не занятые на работе или учебе больные хуже выполнили программу.			
4. У женщин выше уровень занятости.			
5. Не занятые на работе или учебе чаще используют стимулирующие средства.			
5. Больные с незаконченным средним или высшим образованием образованием реже используют успокоительные препараты.			

Распределения, связанные с нормальным

$\xi_1, \xi_2, \ldots, \xi_n$ последовательность независимых стандартно распределенных случайных величин

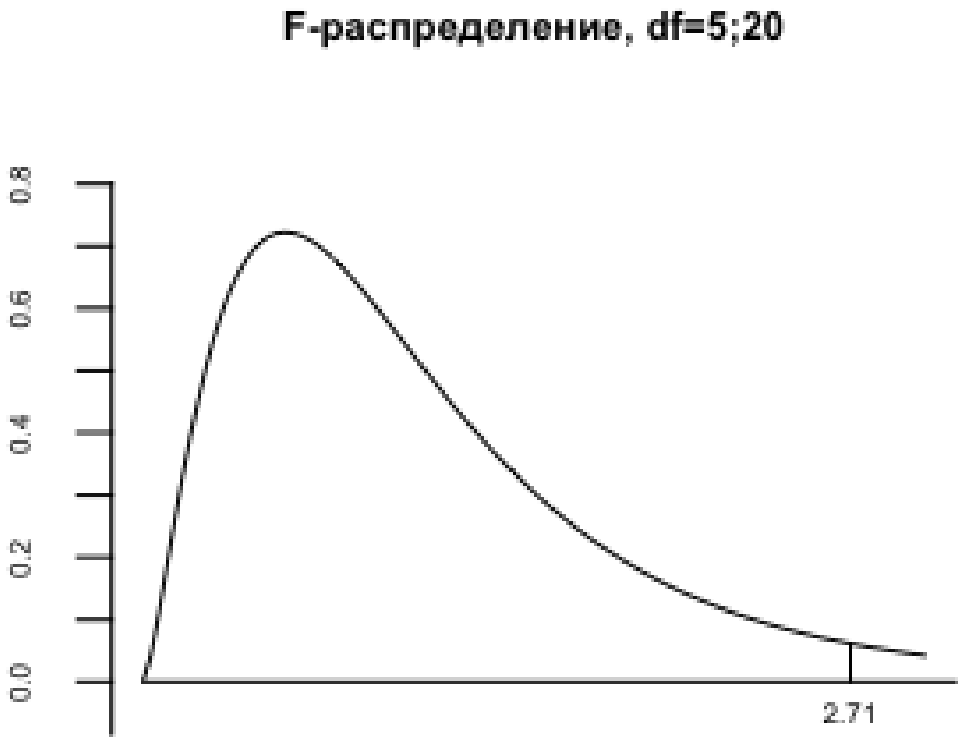
Распределения	хи-квадрат	Стьюдента	Фишера
Степени свободы	n	n	n и m
	$\eta = \sum\limits_{i=1}^{n} \xi_i^2$	$\zeta = \dfrac{\xi_0}{\sqrt{\frac{1}{n}\sum\limits_{k=1}^{n} \xi_k^2}}$	$\vartheta = \dfrac{\frac{1}{n}\sum\limits_{k=1}^{n} \xi_k^2}{\frac{1}{m}\sum\limits_{k=1}^{m} \eta_k^2}$

Квантиль распределения Стьюдента

Например, 0.95-квантиль распределения Стьюдента со степенью свободы 20 равна 1.72, то есть $P\{t < 1.72\} = 0.95$. В R можно использовать функцию $qt(0.95, 20)$. Наоборот $pt(1.72, 20)$ приведет к значению 0.95. В Excel используется функция СТЬЮДРАСП или TDIST.

P-квантили распределения Фишера

P-квантили распределения Фишера с n_1, n_2 степенями свободы ищутся при помощи функции $qf(P, n_1, n_2)$. Если по некоторой x квантили нужно найти соответствующую вероятность, то используется функция $pf(x, n_1, n_2)$, например, функция $pf(2.71, 5, 20)$ приводит к 0.95. В Excel используется функция FPACП или FDIST.



Проверка статистических гипотез

- Для проверки статистической гипотезы H_0 относительно параметров θ или других свойств генеральной совокупности с функцией распределения $F(x|\theta)$ случайной величины ξ используются выборочные наблюдения $x_1, \ldots, x_n$. Альтернативную гипотезу будем обозначать через H_1.
- Гипотеза называется простой, если она однозначно определяет функцию распределения, например, $H_0 : \theta = \theta_0$, иначе сложной, например, $H_0 : \theta > \theta_0$.
- Правило, согласно которому отвергается гипотеза, называется статистическим критерием, а используемая для проверки гипотезы функция от выборочных наблюдений $x_1, \ldots, x_n$ называется статистикой критерия.

Принцип маловероятных событий

- Проверка гипотез основана на принципе маловероятных событий, согласно которому события, вероятность которого мала, считаются невозможными.
- В качестве „малой“ принимают $\alpha = 0.05$ — вероятность выигрыша в игре в кости.
- Гипотезу отвергают тогда, когда наблюдаемое значение статистики попадает в критическую область V, которую выбирает так, чтобы вероятность $\mathrm{P}(\mathrm{V}|\mathrm{H}_0)$ была мала,

$$\mathrm{P}(\mathrm{V}|\mathrm{H}_0) \leq \alpha\,.$$

Построение критерия

- Пусть имеется выборка наблюдений $x_1, \ldots, x_n$ случайной величины, имеющей распределение со средним μ и дисперсией σ^2.
- Статистика – выборочное среднее $\bar{x} = \frac{1}{n} \sum_{j=1}^{n} x_j$.
- при увеличении n выборочное средее $\bar{x}$ имеет нормальное распределение с тем же средним μ и дисперсией σ^2/n.
- Статистика критерия $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$

ПРИМЕР.

На $n = 25$ автомобилях с усовершенствованным двигателем средний расход бензина составил $\bar{x} = 9.3$ л на 100 км. Считая выборку нормальной $\mathcal{N}(\mu, \sigma)$ с дисперсией $\sigma^2 = 4$, выясним, нельзя ли наблюдаемое улучшение технической характеристики считать случайным.

- Проверим гипотезу $H_0 : \mu = 10$ о том, что расход топлива не изменился, В качестве альтернативной рассмотрим гипотезу $H_1 : \mu < 10$.
- Статистика критерия $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$
- Критическая область $V = (-\infty; -1.645]$, $P(V) = 0.05$.
- Значение статистики $Z = -1.75 \in V$, следовательно, гипотеза H_0 отвергается в пользу альтернативной H_1, уменьшение топлива значимо и не может быть объяснено случайностью.

Доверительный уровень вероятности или значимость, p-значение

По заданному значению x_* статистики X считается доверительный уровень вероятности или p-значение.

p-значение в случае односторонней альтернативы $H_1 : \Theta > \Theta_0$

равно вероятности получить случайное значение больше наблюдаемого, $p = P\{X > x_*\}$.

p-значение в случае односторонней альтернативы $H_1 : \Theta < \Theta_0$

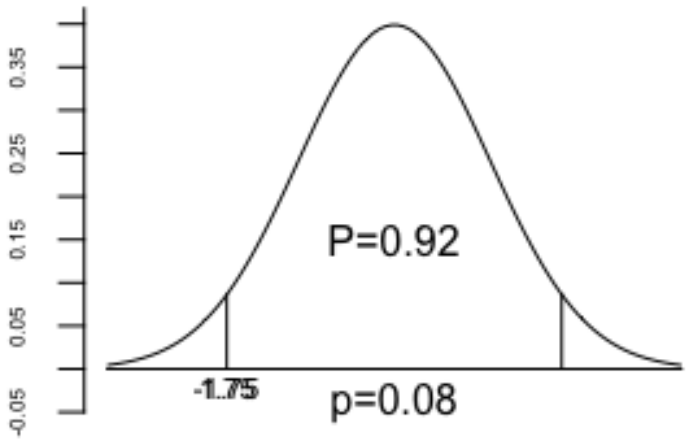
равно вероятности получить случайное значение меньше наблюдаемого, $p = P\{X < x_*\}$.

Нулевая гипотеза отвергается в пользу альтернативной в случае, когда p меньше заданного уровня значимости α, то есть при $p < \alpha$. В качестве уровня значимости обычно берется значение 0.05.

p-значение в случае односторонней альтернативы $H_1 : \Theta \neq \Theta_0$

равно вероятности получить случайное значение больше наблюдаемого по абсолютной величине, $p = P\{|X| > |x_*|\}$.

В примере при альтернативе $H_1 : \mu \neq 10$ значимость равна $p = 0.08$, следовательно, при $\alpha = 0.1$ нулевая гипотеза $H_0 : \mu = 10$ отвергается, а при $\alpha = 0.05$ нет оснований отвергнуть $H_0 : \mu = 10$.



Ошибки первого и второго рода, мощность, p-значение

- Ошибка первого рода возникает, если отвергается верная гипотеза, $\alpha = \mathrm{P}(\mathrm{V}|\mathrm{H}_0)$,
- ошибка второго рода, если не отвергается ложная гипотеза $\beta = \mathrm{P}(\bar{\mathrm{V}}|\mathrm{H}_1)$.
- Мощностью называется вероятность $\mathrm{P}(\mathrm{V}|\mathrm{H}_1) = 1 - \beta$.

Критерий для проверки H_0 называется несмещенным, если для любой альтернативы вероятность попадания в критическую область будет больше, чем вероятность попадания в нее при справедливости нулевой гипотезы, то есть $\forall \mathrm{H}_1 \ \mathrm{P}(\mathrm{V}|\mathrm{H}_0) \leq \mathrm{P}(\mathrm{V}|\mathrm{H}_1)$.

Пример вычисления ошибки первого рода и мощности критерия

Через $\Phi(\mathrm{x})$ обозначена функция стандартного нормального распределения. Для проверки гипотезы $\mathrm{H}_0 : \mu = 10$, при $\sigma = 2$, $\mathrm{n} = 25$ с критической областью V, соотвествующей неравенству $\bar{\mathrm{x}} < \mu - \mathrm{Z}_{0.95}\frac{\sigma}{\sqrt{\mathrm{n}}} = 9.342$, где $\mathrm{Z}_{0.95}$ — квантиль $\mathcal{N}(0,1)$, ошибка первого рода равна

$$\alpha = \mathrm{P}\{\bar{\mathrm{x}} < 9.342 | \mu = 10\} = \mathrm{P}\left\{ \frac{\bar{\mathrm{x}} - \mu}{\sigma/\sqrt{\mathrm{n}}} < \frac{9.342 - \mu}{\sigma/\sqrt{\mathrm{n}}} \;\middle|\; \mu = 10 \right\} = \\ = \Phi(-1.645) = 0.05\,.$$

Если $\mathrm{H}_1 : \mu = 9$, тогда $1 - \beta = \mathrm{P}\{\bar{\mathrm{x}} < 9.342 | \mu = 9\} =$

$$= \mathrm{P}\left\{ \frac{\bar{\mathrm{x}} - \mu}{\sigma/\sqrt{\mathrm{n}}} < \frac{9.342 - \mu}{\sigma/\sqrt{\mathrm{n}}} | \mu = 9 \right\} = \Phi(0.855) = 0.804\,.$$

Таким образом, мощность критерия $\bar{\mathrm{x}} < 9.342$ равна 0.804.

Усовершенствование будет установлено в 80% случаев.

Статистика для проверка гипотезы $H_0 : \mu = \mu_0$ при неизвестной дисперсии

Выборка $x_1, \ldots, x_n$ со средним $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ и выборочной дисперсией

$$S^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2 = \frac{1}{n-1}\left(\sum_{j=1}^{n} x_j^2 - n\bar{x}^2\right).$$

Теорема о статистике Стьюдента

Если $x_1, \ldots, x_n \sim \mathcal{N}(\mu, \sigma)$, то 1) $\bar{x}$ и S^2 не зависимы,
2) $\bar{x} \sim \mathcal{N}(\mu, \frac{\sigma}{\sqrt{n}})$, 3) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$,
4) $T = \frac{\bar{x}-\mu}{S}\sqrt{n} \sim T(n-1)$.

Пусть $\xi_0 = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$, $\quad \eta = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, тогда

$$T = \frac{\xi_0}{\sqrt{\frac{1}{n-1}\eta}} = \frac{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{1}{n-1}\frac{(n-1)S^2}{\sigma^2}}} = \frac{\bar{x}-\mu}{S}\sqrt{n} \sim T(n-1).$$

Задача на вычисление ошибки первого рода и мощности при неизвестной дисперсии

Попробуем поручить компьютеру отобрать детей для спортивной школы. Проанализируем данные $x_1, \ldots, x_n$, $n = 25$, какого-то норматива со средним $\bar{x} = 9.3$ и выборочной дисперсией $S^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2 = 4$.

Проверим гипотезу $H_0 : \mu = 10$ против альтернативы $H_1 : \mu < 10$.

Статистика $T = \frac{\bar{x}-\mu}{S}\sqrt{n} = \frac{9.3-10}{2/5} = -1.75$.

p-значение $P\{T < -1.75\} = 0.046 < 0.05$, поэтому отклонение выборочного среднего $\bar{x} = 9.3$ от 10 является значимым.

Проверка

Вопросы по проверке независимости качественных признаков	да-нет	P	зн.-незн.
1. Прием успокоительных средств является препятствием для выполнения программы.	30, 18	0.074	
2. Больные, принимающие вызывающие галлюцинации препараты, хуже выполнили программу.	30, 17	0.029	
3. Не занятые на работе или учебе больные хуже выполнили программу.	25, 36	0.07	
4. У женщин выше уровень занятости.	29, 23	0.38	
5. Не занятые на работе или учебе чаще используют стимулирующие средства.	15, 5	0.03	
5. Больные с незаконченным средним или высшим образованием образованием реже используют успокоительные препараты.	5, 25, 14, 56	0.0124	