SkillFactory

Тестирование гипотез

Statistics means never having to say you are certain

Это способ принятия решений

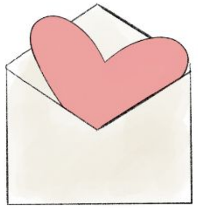
- Мы живем в мире неопределенности.
- Поэтому, принимая решения, мы не можем избежать ошибок.
- Тестирование гипотез – инструмент принятия оптимальных решений в условиях неполной информации.

Тестирование: пример

1. Сформулировать гипотезы

   H_o: «Ваня любит меня» **(нулевая гипотеза)**

   H_A: «Ваня не любит меня» **(альтернативная гипотеза)**



2. Собрать наблюдения.

3. Мысленный эксперимент: Если H_o верна, как должен вести себя Ваня?

4. Вероятность **p-value:** насколько ожидаемо наблюдаемое поведение Вани, предполагая гипотезу H_o?

5. Вывод. Если p-value мала (< **уровня значимости α**) H_o отвергается в пользу H_A. Если вероятность не мала – не отвергается.

=> Таким образом, Маша решила, Ваня её НЕ любит.

«Но это не точно!»

Ошибки тестирования

- **Ошибка I рода**

Верна H_o, но ее отвергли

- **Ошибка II рода**

Верна H_A, но H_o не отвергли

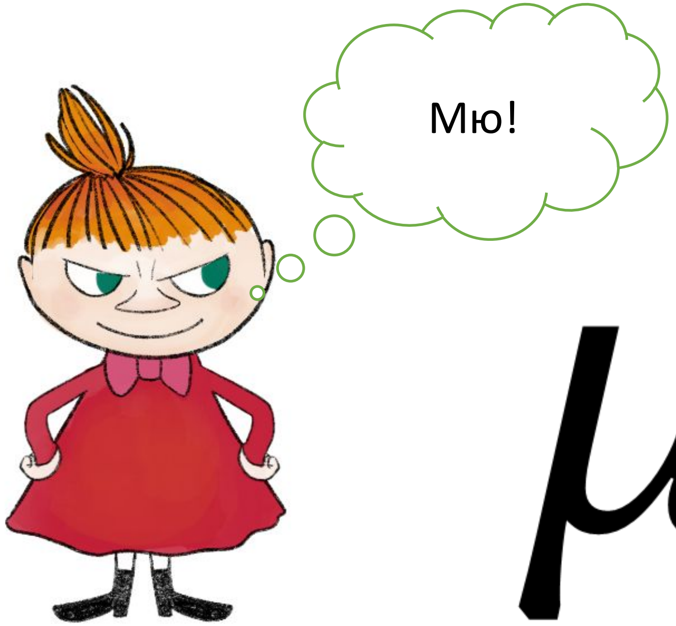
$$\text{P(ошибка } I\text{)} = \alpha$$

$$\text{P(ошибка } II\text{)} = \beta$$

Крошка Мю

-

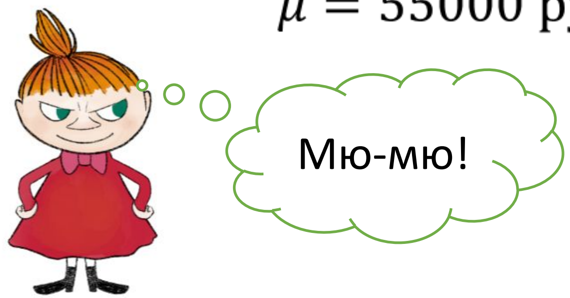
Тестирование гипотез об истинном среднем значении μ (‘мю’)



$$\mu$$

Пример 1. Средняя зарплата

$\mu = 55000$ рублей?



X – заплата жителя города

$\sigma = 15500$

Выборка: $n = 40$, $\bar{X} = 35000$

- H_o: $\mu = 55000$
- H_A: $\mu < 55000$

Пример 1. Средняя зарплата

Шаг 1. H_o: $\mu = 55000$

H_A: $\mu < 55000$

Шаг 2. $n = 40, \bar{X} = 35000,\ \sigma = 15500$

$\overline{\mathrm{X}} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

Шаг 3. $\overline{\mathrm{X}} \approx N(55000, \frac{15500}{\sqrt{40}})$

Шаг 4. $z_{st} = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{35000-55000}{\frac{15500}{\sqrt{40}}} = -8.16$

На 8 станд. откл. ниже, чем ожидали по H_o!

Пример 1. Средняя зарплата

- $$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \approx N(0,1)$$

Шаг 4. Найти **p-value** – вероятность получить такую или более “странную” выборку при условии, что нулевая гипотеза верна.

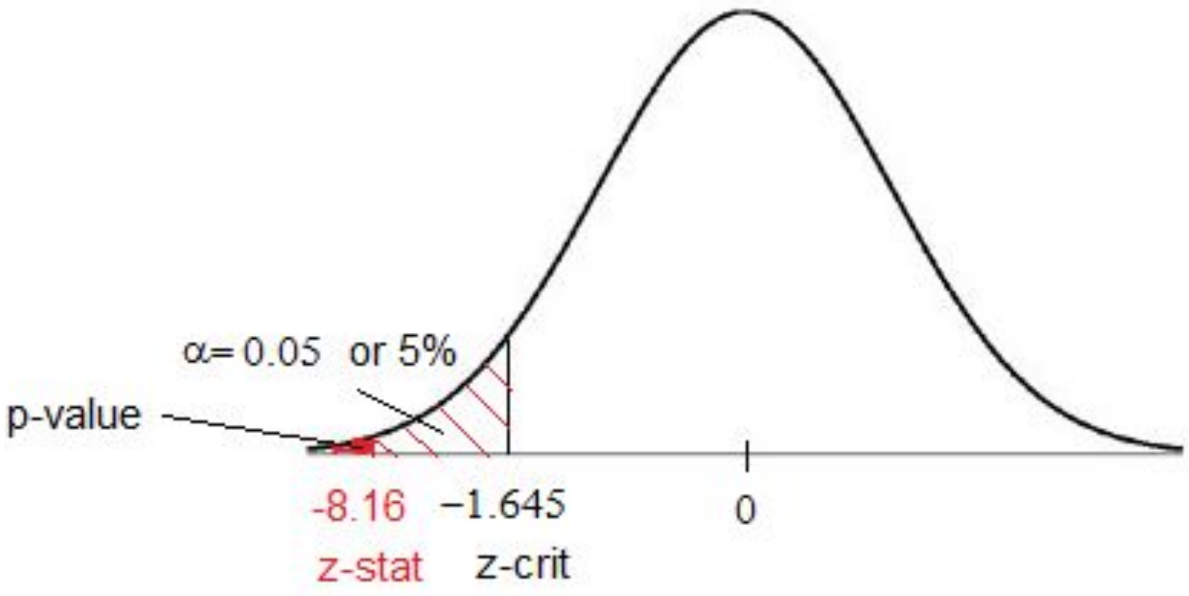
$$\mathrm{p - value} = \mathrm{P}(\bar{\mathrm{X}} \leq 35000 | H_0) = P(Z \leq -8.16) = 1.7 \cdot 10^{-16}$$

Шаг 5. Принять решение. $\alpha = 0.05$
$p - value = 0.00000000000000017 < \alpha$

Таким образом, мы отвергли H_o в пользу H_A

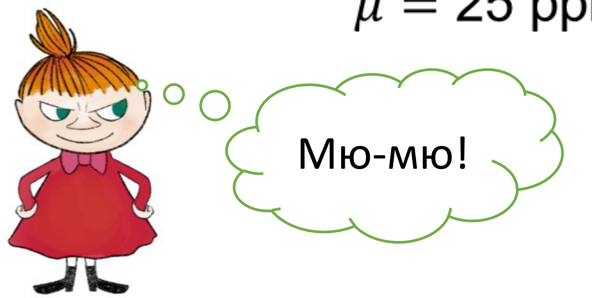
Пример 1. Средняя зарплата

На доске



Пример 2. Уровень загрязнения

$\mu = 25$ ppm?



X – содержание вредных в-в в атмосфере

$\sigma = 14$ ppm, $X \sim N$

Выборка: $n = 20$, $\bar{X} = 37$

- H_0: $\mu = 25$
- H_A: $\mu. > 25$

Пример 2. Уровень загрязнения

Шаг 1. $H_0: \mu = 25$

$H_A: \mu > 25$

Шаг 2. : $n = 20, \ \overline{X} = 37, \ \sigma = 14\, ppm, \ X \sim N$

$\overline{X} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

Шаг 3. По H_0 $\overline{X} \approx N(25, \frac{14}{\sqrt{20}})$

Шаг 4. $z_{st} = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{37-25}{\frac{14}{\sqrt{20}}} = 3.83$

$p - \text{value} = P(\overline{X} \geq 37 | H_0) =$

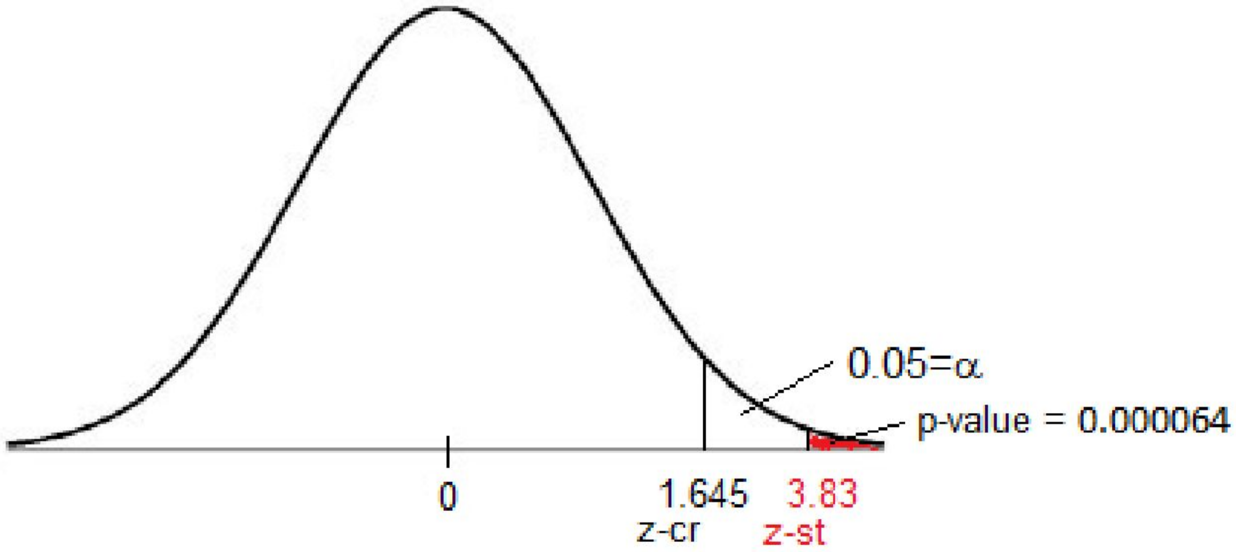
$= P(Z \geq 3.83) = 0.000064$

Шаг 5. $p - \text{value} = 0.000064 < \alpha = 0.05$

Таким образом, мы отвергли H_0 в пользу H_A

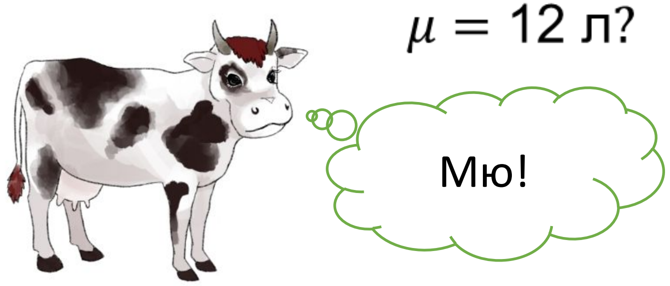
Пример 2. Уровень загрязнения

На доске



Пример 3. Удой молока

$\mu = 12$ л?



X – суточный удой молока с коровы

$\sigma = 5$ л, $X \sim N$

Выборка: $n = 25$, $\overline{X} = 10.8$ л

- H_0: $\mu = 12$
- H_A: $\mu \neq 12$

Пример 3. Удой молока

Шаг 1. H_0: $\mu = 12$

H_A: $\mu \neq 12$

Шаг 2: $n = 25,\ \overline{X} = 10.8,\ \sigma = 5,\ X \sim N$

$\overline{X} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

Шаг 3. По H_0 $\overline{X} \approx N(12, \frac{5}{\sqrt{25}})$

Шаг 4. $z_{st} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{10.8 - 12}{\frac{5}{\sqrt{25}}} = -1.2$

$p - \text{value} = P(|Z| \geq 1.2) =$

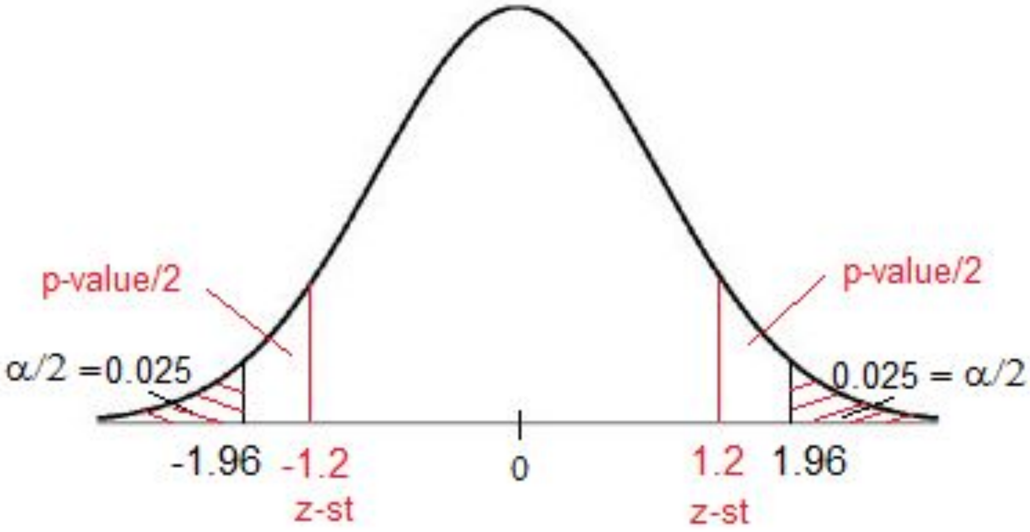
$= 2P(Z \leq -1.2) = 2 \cdot 0.115 = 0.23$

Шаг 5. $p - \text{value} = 0.23 > \alpha = 0.05$

Таким образом, мы не отвергли H_0

Пример 3. Удой молока

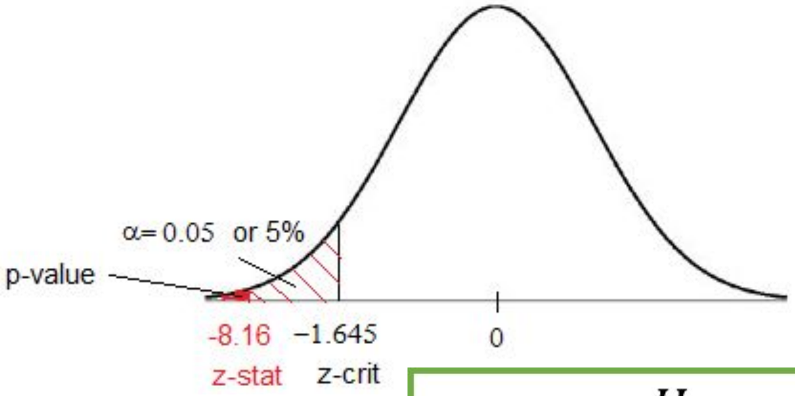
На доске

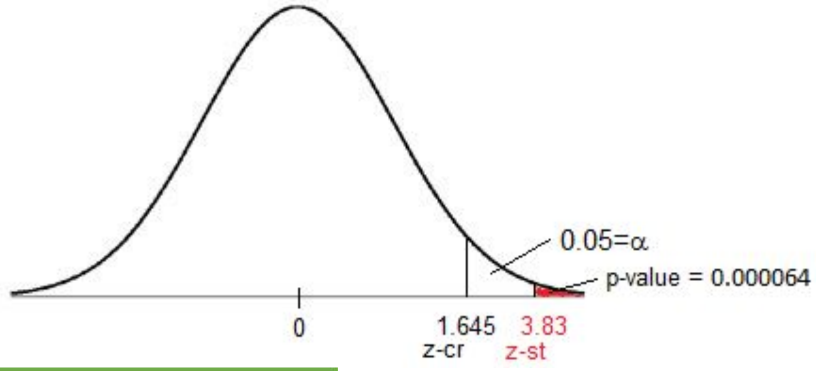


Как считать p-value?

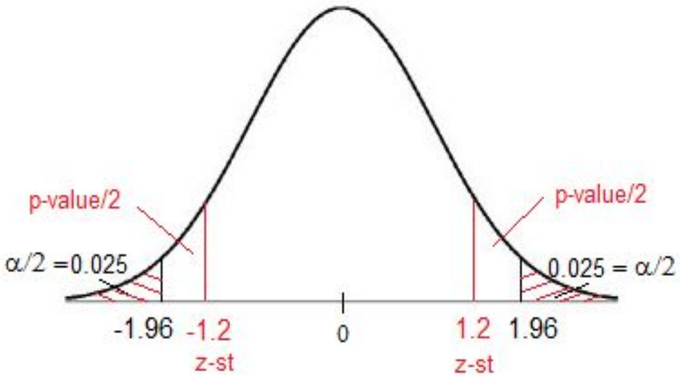
- $H_A\colon \mu < 55000$
  $\text{p} - \text{value} = P(Z \le -8.16)$

- $H_A\colon \mu > 25$
  $\text{p} - \text{value} = P(Z \ge 3.83)$





- $H_A\colon \mu \neq 12$
  $\text{p} - \text{value} = 2P(Z \le -1.2)$



Тестирование: порядок действий

Шаг 1. Сформулировать гипотезы:

$\mathrm{H_0}$: …

$\mathrm{H_A}$: …

Шаг 2: Собрать наблюдения, посчитать выборочные статистики: $\mathrm{n} = \ldots$ и т.д.

Шаг 3. Мысленный эксперимент, ожидаемое распределение

Шаг 4. Посчитать z_{st} и $\mathrm{p-value}$

Шаг 5. Сравнить $\mathrm{p-value}$ с α и сделать вывод

Тестовая статистика = $\frac{\text{Оценка}-\text{Параметр}_{\mathbf{0}}}{\boldsymbol{s}_{\text{оценки}}}$

Тестирование гипотез о среднем μ

Тестовая статистика для μ

?

$$z_{st} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Если σ известно

$$t_{st}(k) = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

k=n-1

Если σ неизвестно

Тестирование гипотез о пропорции p

- $\hat{p} \approx N\left(p, \sqrt{\frac{p(1-p)}{n}}\right) \Rightarrow \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \approx Z \sim N(0,1)$

$$z_{st} = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Тестовая статистика = $\frac{\text{Оценка} - \text{Параметр}_{\mathbf{0}}}{\boldsymbol{s}_{\text{оценки}}}$

Допущения:

1. Выборка взята случайно
2. Выборка достаточно велика

$(n\hat{p} \geq 5,\ n(1-\hat{p}) \geq 5)$

Пример. Тестирование гипотез о p

Группа медиков утверждает, что пропорция граждан, подверженных опасному заболеванию, составляет 35% населения и требует открытия гос.программы по предотвращению заболевания. Минздрав выделит бюджет, если будет доказано, что пропорция действительно не меньше 35%. Тест должен быть сделан на 5% уровне значимости.

В случайной выборке из 185 человек 77 человек оказались в группе риска данного заболевания.

Будет ли Минздрав выделять средства на программу по борьбе с ним?

p – пропорция граждан, подверженных болезни

$$\hat{p} = \frac{77}{185} \approx 0.42, n = 185$$

Пример. Тестирование гипотез о $\boldsymbol{p}$

Шаг 1. $H_0: p = 0.35$

$H_A: p > 0.35$

Шаг 2. $\hat{p} = \frac{77}{185} \approx 0.42, n = 185$

Шаг 3. Выборка случайная,

$n\hat{p} = 77 > 5$, $n(1-\hat{p}) = 108 > 5$. Значит, $\hat{p} \approx N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$

По $H_0: \hat{p} \approx N\left(p_0, \sqrt{\frac{p_0(1-p_0)}{n}}\right)$

Шаг 4. $z_{st} = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.42-0.35}{\sqrt{\frac{0.35\,(1-0.35)}{185}}} \approx 1.89$

$p - \text{value} = P(Z \geq 1.89) \approx 0.029$

Шаг 5. $p - \text{value} \approx 0.029 < \alpha = 0.05,$

H_0 отвергнута в пользу $H_A: p > 0.35$

Минздрав выделит средства на программу

Разность средних - распределение

- *(На доске)*

$$\overline{X_1} \sim N, \overline{X_2} \sim N$$

$$\overline{X_1} - \overline{X_2} \sim N(?\,,?\,)$$

$$E(\overline{X_1} - \overline{X_2}) = E(\overline{X_1}) - E(\overline{X_2}) = \mu_1 - \mu_2$$

$$Var(\overline{X_1} - \overline{X_2}) = Var(\overline{X_1}) + Var(\overline{X_2}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$\overline{X_1} - \overline{X_2} \sim N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

Разность средних - тестирование

- *(На доске)*

$$\overline{X_1} - \overline{X_2} \sim N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

H_0: $\mu_1 = \mu_2$ $(\mu_1 - \mu_2 = 0)$
H_A: $\mu_1 > \mu_2$ ИЛИ $\mu_1 < \mu_2$ ИЛИ $\mu_1 \neq \mu_2$

$$\text{По } H_0: \overline{X_1} - \overline{X_2} \sim N(0, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

$$z_{st} = \frac{\overline{x_1} - \overline{x_2} - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \; t_{st}(k) = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Если $\sigma_1 = \sigma_2 = \sigma$

$$z_{st} = \frac{\overline{x_1} - \overline{x_2}}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \; t_{st}(m) = \frac{\overline{x_1} - \overline{x_2}}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Тестирование гипотез о $\mu_1 - \mu_2$

$$z_{st} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

$$t_{st}(k) = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

$$t_{st}(m) = \frac{\overline{x_1} - \overline{x_2}}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

σ_1, σ_2 известны

σ_1, σ_2 неизвестны
$\sigma_1 \neq \sigma_2$

σ_1, σ_2 неизвестны,
$\sigma_1 = \sigma_2$

Допущения:

1. Выборки взяты случайно
2. Выборки взяты независимо друг от друга
3. Выборки достаточно велики* ($n_1 \geq 30, n_2 \geq 30$)

* *Или $X_1 \sim N$ and $X_2 \sim N$*

Проведение эксперимента

- Экспериментальная и контрольная группы:

Экспериментальная группа – получает лекарство;

Контрольная группа – не получает лекарство (получает «пустышку» – плацебо).

Эффект Плацебо – эффект самореализующихся ожиданий испытуемого.

Плацебо – «пустышка», выдаваемая в контрольной группе, чтобы эффект плацебо был одинаков в обеих группах.

Пример. Тестирование гипотез о $\mu_1 - \mu_2$

исследование по снижению уровня холестерина среди мужчин. 20 мужчин с повышенным уровнем холестерина выбраны случайно из соответствующей группы пациентов крупной клиники. Далее они были случайно разделены на группы: 10 мужчин попали в группу А, пациентам которой давали плацебо. Остальные 10 пациентов попали в группу В, где пациенты получали лекарство. Через 3 месяца все был измерен уровень холестерина у обеих групп пациентов.

В группе А среднее снижение уровня холестерина составило 10.2 мл/дл со стандартным отклонением 7.66 мл/дл. В группе В среднее снижение составило 16.4 мл/дл со стандартным отклонением 9.4 мл/дл.

Известно, что динамика в уровне холестерина среди пациентов имеет приблизительно нормальное распределение.

Протестируйте гипотезу об эффективности лекарства на 1% уровне значимости

X_A, X_B - снижение ур-ня холестерина в группах А,В. X_A, $X_B \approx N$

$\bar{x}_A = 10.2,\ \bar{x}_B = 16.4,\ s_A = 7.66,\ s_B = 9.4,\ n_A = n_B = 10,\ \alpha = 0.01$

Пример. Тестирование гипотез о $\boldsymbol{\mu_1 - \mu_2}$

Шаг 1. $H_0\colon \mu_A = \mu_B$ (или $\mu_A - \mu_B = 0$)

$H_A\colon \mu_A < \mu_B$

Шаг 2. $\bar{x}_A = 10.2,\ \bar{x}_B = 16.4, s_A$=7.66, s_B=9.4, $n_A = n_B = 10$

Шаг 3. $X_A, X_B \approx N$, выборки случайные и независимые.

$\overline{X_1} - \overline{X_2} \sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$

Шаг 4. $t_{st}(\boldsymbol{k}) = \frac{\overline{\boldsymbol{x}_A} - \overline{\boldsymbol{x}_B}}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} = \frac{10.2 - 16.4}{\sqrt{\frac{7.66^2}{10} + \frac{9.4^2}{10}}} \approx -1.617, \mathrm{k} \approx 17$

p-value= $P(t(17) < -1.617) \approx 0.062$

Шаг 5. p-value> α=0.01,

H_0 не отвергнута на 1% уровне значимости

Распределение разности пропорций

- $\widehat{p_1} \approx N\left(p_1, \sqrt{\frac{p_1(1-p_1)}{n_1}}\right), \widehat{p_2} \approx N\left(p_2, \sqrt{\frac{p_2(1-p_2)}{n_2}}\right)$

$$\hat{p}_1 - \hat{p}_2 \sim N(?\,,?)$$

$E(\hat{p}_1 - \hat{p}_2) = E(\hat{p}_1) - E(\hat{p}_2) = p_1 - p_2$

$Var(\hat{p}_1 - \hat{p}_2) = Var(\hat{p}_1) + Var(\hat{p}_2)$

$= \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$

$\hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}})$

H_0: $p_1 = p_2$ $(p_1 - p_2 = 0)$

$p_1 = p_2 = p$

По H_0: $\hat{p}_1 - \hat{p}_2 \sim N(0, \sqrt{p(1-p) \cdot \frac{1}{n_1} + \frac{1}{n_2}})$

$$z_{st} = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{p(1-p) \cdot \frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\hat{p} = \frac{m_1 + m_2}{n_1 + n_2} = \frac{\hat{p}_1.n_1 + \hat{p}_2.n_2}{n_1 + n_2}$$

Распределение разности пропорций

•

$$z_{st} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p}) \cdot \frac{1}{n_1} + \frac{1}{n_2}}}$$

Допущения:

1. Выборки взяты случайно
2. Выборки независимы друг от друга
3. Выборки достаточно велики

$n_1\widehat{p_1} \geq 5,\ n_1(1-\widehat{p_1}) \geq 5,$

$n_2\widehat{p_2} \geq 5,\ n_2(1-\widehat{p_2}) \geq 5$

Пример. Тестирование гипотез о $\boldsymbol{p_1 - p_2}$

За 2 месяца до выборов опрос показал, что 460 из 800 опрошенных поддерживают определенного кандидата. Однако за неделю до выборов другой независимый опрос показал поддержку в 520 из 1000 опрошенных.

Протестируйте на 10% уровне значимости падение поддержки кандидата за рассматриваемый период.

p_1, p_2– поддержка кандидата за 2 месяца и за неделю до выборов.

$\hat{p}_1 = \frac{460}{800} = 0.575$, $\hat{p}_2 = \frac{520}{1000} = 0.520$, $n_1 = 800$, $n_2 = 1000$

Пример. Тестирование гипотез о $p_1 - p_2$

Шаг 1. $H_0: p_1 = p_2$ (или $p_1 - p_2 = 0$)

$H_a: p_1 > p_2$

Шаг 2. $\hat{p}_1 = \frac{460}{800} = 0.575,\ \hat{p}_2 = \frac{520}{1000} = 0.520,$

$n_1 = 800,\ n_2 = 1000$

Шаг 3. Выборки случайные, независимые и:

$n_1\widehat{p_1} = 460 > 5,\ n_1(1 - \widehat{p_1}) = 340 > 5,$

$n_2\widehat{p_2} = 520 > 5,\ n_2(1 - \widehat{p_2}) = 480 > 5$

Значит, $\hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \sqrt{\frac{p_1 \cdot (1-p_1)}{n_1} + \frac{p \cdot (1-p_2)}{n_2}})$

Шаг 4. $\hat{p} = \frac{\hat{p}_1 . n_1 + \hat{p}_2 . n_2}{n_1 + n_2} = \frac{460+520}{800+1000} = 0.544$

$$z_{st} = \frac{p_1 - p_2}{\sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{0.575 - 0.520}{0.0236} = \frac{0.055}{0.0236} = 2.33$$

p-value= $P(Z > 2.33) = 0.0099 \approx 0.01$

Шаг 5. p-value$\approx 0.01 < \alpha = 0.1$, H_0 отвергается.

Значит, уровень поддержки упал.