Лабораторная работа №2

Описание данных. Статистический вывод

Цель: Знакомство с этапом понимание данных стандарта CRISP-DM.

Задача этапа – найти, описать основные закономерности, которые содержатся в данных и статистически их подтвердить, попытаться выявить связи между этими данными, фактами, знаниями, получить что-то новое, понять, что они могут дать.

Задание

Получив первоначальное представление о данных, рассмотрите закономерности, присущие данным.

Для задачи, которую вы сформулировали в предыдущей работе:

Подумайте, какие статистические гипотезы могут подтвердить ил опровергнуть ваши предположения.

Сформулируйте гипотезы о равенстве выборочных средних и долей категориальных, разделяя переменные. Используйте закономерности, особенности, которые вы смогли обнаружить в предыдущей работе. Проследите правильность формулировок гипотез H_0 и H_1 . Проведите испытание, используя соответствующие критерии.

Для того, чтобы сформулировать полезные гипотезы:

Подумайте, что вы хотите выяснить, проверить.

Можно использовать фильтры разного уровня. Например, чтобы сравнить учебные успехи студентов, имеющих двоих и более друзей и тех, у кого друзей меньше – отфильтруйте по этому признаку показатели успеваемости, посчитайте среднее и доверетельные интервалы. Если доверительные интервалы пересекаются – сформулируйте гипотезы. Результат позволит сделать заключение – отличается ли успеваемость студентов, которые общаются мало и их более общительных товарищей

В отчете должны быть приведены: вид используемого теста (одновыборочный , двухвыборочный, ранговый и т.д.), формулировки нулевой и альтернативной гипотез, вычисленные и взятые из таблиц статистики

Теория

СТАТИСТИЧЕСКИЕ ИСПЫТАНИЯ НА НЕЗАВИСИМОСТЬ

Статистическое испытание гипотез проводится, когда по ограниченной информации (выборке) нужно определить, могут ли определенные предположения быть приняты. Например,

- Время между отказами системы составляет в среднем три месяца.
- Имеется существенное различие между средним уровнем задержки поставок от двух поставщиков.
- Новая система планирования уменьшила затраты.

Рассмотрим пример с расфасовкой сахарного песка.

Фасовочная машина должна быть установлена так, чтобы при расфасовке каждый пакет в среднем имел определенный вес. С момента установки машины средняя наполняемость пакета вряд ли останется постоянной — отклонения от нормы будут возрастать по мере износа машины.

Производитель может организовать контроль, в котором берется выборка, вычисляются $\overline{x}$, s и доверительный интервал для веса пакета μ в генеральной совокупности. Если доверительный интервал «накрывает» требуемое значение веса, значит повода для перенастройки машины пока нет.

Можно использовать другую процедуру – испытание гипотез для оценки того, правильно ли работает машина.

Чтобы выполнить испытание, требуется две гипотезы: нулевая гипотеза, обозначаемая H_0 и альтернативная – H_1.

Пусть машина была настроена на вес 550 г. Производитель может полагать, что машина все еще настроена правильно, и средний вес пакетов по-прежнему составляет 550 г. Его предположение составляет нулевую гипотезу и может быть записано как

$$H_0: \mu=550$$

Нулевая гипотеза всегда формулируется для утверждения того, что выборочная статистика согласуется с принятым параметром генеральной совокупности.

Эта гипотеза будет принята, если не найдется достаточно веских доказательств, свидетельствующих против нее.

Могут быть выдвинуты три варианта альтернативной гипотезы:

$$H_1: \mu>550$$

$$H_1: \mu<550$$

$$H_1: \mu\neq550$$

Первая гипотеза предполагает, что средний вес пакетов сахара существенно больше, чем 550 г (машина «пересыпает»). Вторая гипотеза предполагает, что средний вес пакетов сахара существенно меньше, чем 550 г (машина «недосыпает»). Третья гипотеза предполагает, что вес пакетов существенно отличен от 550 г.

Из трех вариантов нужно выбрать один, исходя из условий реальной ситуации.

Первые две гипотезы требуют *односторонних* испытаний, а третья гипотеза – *двустороннего* испытания. Такие испытания проводятся на основе определенных правил и базируются на суждении лица, принимающего решения.

Какой бы результат ни был получен, мы никогда не можем определенно (на 100%) доказать или опровергнуть нулевую гипотезу. Все, что мы можем сделать, — это или признать то, что нулевая гипотеза почти наверняка верна, или, что правильность нулевой гипотезы маловероятна. Уверенность на 100% достижима только тогда, когда измерена вся генеральная совокупность. Поэтому может случиться, что нулевая гипотеза будет отклонена, хотя она была истинной.

ИСПЫТАНИЕ ГИПОТЕЗЫ: ВОЗМОЖНЫЕ РЕЗУЛЬТАТЫ

Имеются четыре возможных результата при испытании гипотезы:

- Вы примете гипотезу, истинность которой покажет испытание, и она действительно была истинной (т.е. Вы пришли к правильному заключению).
- Вы отклоните гипотезу, ложность которой покажет испытание, и она действительно была ложной (т.е. Вы снова пришли к правильному заключению).
- Вы отклоните гипотезу, т.к. испытание покажет ее ложность, но она в действительности была истинной (т.е. Вы пришли к неправильному заключению).

Такая ошибка относится к ТИПУ I и связана с уровнем значимости α.

- Вы примете гипотезу, т.к. испытание покажет ее истинность, но она в действительности была ложной (т.е. Вы снова пришли к неправильному заключению). Такая ошибка относится к ТИПУ II.

Вероятность отклонения правильной нулевой гипотезы, т.е. ошибки ТИПА I , должна задаваться до проведения испытания. Эта величина называется уровнем значимости α и представляет собой потенциальную ошибочность любых сделанных впоследствии выводов. Чем меньше уровень значимости, тем меньше потенциальная ошибка.

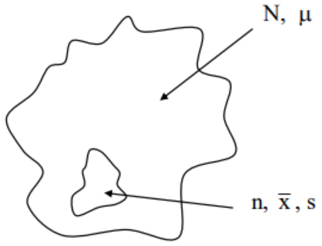
Если Вы выбираете уровень значимости α=0.05, то другими словами Вы можете быть на 95% уверены в правильности принятого решения.

ИСПЫТАНИЕ ГИПОТЕЗЫ О СРЕДНЕМ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

Существуют три типичные ситуации

Испытание гипотезы по выборочному среднему – одна выборка

Известно, каким должно быть среднее значение генеральной совокупности. Взяв выборку, нужно проверить, не изменилось ли оно.



Проверочная статистика $t = \dfrac{(\bar{x}-\mu)}{\dfrac{s}{\sqrt{n}}}$

Число степеней свободы (n-1)

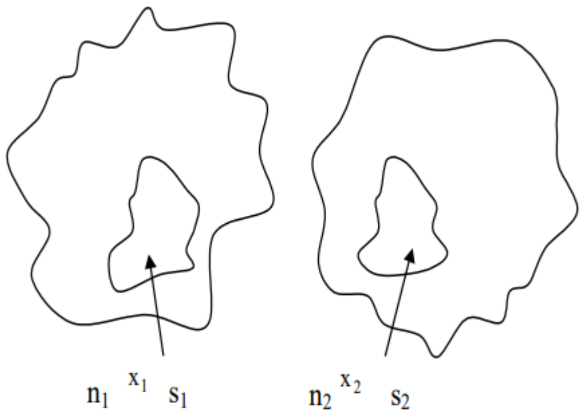
Условие: если выборка маленькая (n<30), то распределение предполагается нормальным

Примеры задач

- Пресс делает отверстия 4 мм в металлических шайбах. Сохраняется ли правильная настройка?
- Известно, что средняя продолжительность рабочей операции составляет 50 минут. На экспериментальном участке используют новый метод. Можно ли утверждать, что он быстрее?
- Компания закупает комплектующие со средним сроком службы 30 дней и имеет возможность сменить поставщика на более качественного. Взята выборка для испытаний. Можно ли полагать, что срок службы будет больше?

Испытание гипотезы по выборочным средним – независимые выборки

Имеются две выборки одинакового или разного размера, элементы которых не связаны между собой. Нужно проверить, существует ли различие между средними значениями совокупностей, из которых выборки взяты.



Проверочная статистика $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$,

где $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1+n_2-2)}$

Число степеней свободы (n_1+n_2–2)

Условие: если выборки маленькие (n_1 <30 или n_2<30), то распределение предполагается нормальным

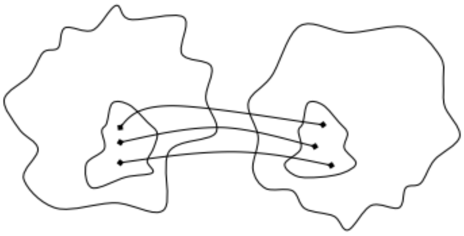
Дополнительное условие: $s_1 \approx s_2$

Примеры задач.

- Сравниваются два типа двигателей по количеству поломок. Есть ли основания утверждать, что один тип ломается чаще?
- Группа рабочих прошла обучение безопасным методам труда. Ведется наблюдение за этой группой и за случайно выбранной контрольной группой. Есть ли основания полагать, что рабочие, прошедшие обучение, реже нарушают правила?

Испытание гипотезы по парным данным – зависимые выборки

Имеются две выборки одинакового размера, элементы в которых составляют пары. Нужно проверить, существует ли различие между средними значениями совокупностей, из которых выборки взяты



.

Проверочная статистика $t = \dfrac{(\bar{x}_d - \mu_d)}{\dfrac{s_d}{\sqrt{n}}}$

Число степеней свободы (n-1)

Условие: если выборка маленькая (n<30), то распределение предполагается нормальным

Примеры задач

- Внесено рационализаторское предложение по ускорению технологической операции. Случайным образом отобрана группа рабочих, и измерено время выполнения каждым из них операции старым способом. Проведено обучение новому способу, и вновь измерено время выполнения операции теми же рабочими. Можно ли считать, что продолжительность операции сократилась?
- Сравнивается качество шин двух производителей. На задние колеса машин разных моделей с разными водителями случайным образом установлена одна шина первого и одно шина второго производителя. После испытания измерен износ. Можно ли считать, что качество шин двух производителей различно?
- Руководство подразделения вводит новую систему показателей для оценки работы. Проведено анкетирование случайно отобранных сотрудников, в котором их просили оценить по балльной шкале «справедливость в оценке труда» до и после изменения. Можно ли полагать, что персонал подразделения считает новую систему более справедливой?

Порядок испытания гипотез следующий:

- Определите, к какому типу относится задача
- Сформулируйте подходящие нулевую и альтернативную гипотезы
- Задайте уровень значимости
- Вычислите *проверочную статистику*.
- Найдите в таблице критическое значение проверочной статистики и сравните его с *модулем* вычисленного значения.

Критическое значение зависит от того, какое испытание проводится – одностороннее или двустороннее, от числа степеней свободы, а также от уровня значимости, который был выбран.

- По результатам сравнения сделайте вывод о необходимости принять или отвергнуть нулевую гипотезу.

Если |проверочная статистика|> |критического значения|, нулевая гипотеза H_0 отклоняется

Если |проверочная статистика|< |критического значения|, нулевая гипотеза H_0 принимается

- Используйте результаты испытания гипотезы, чтобы принять решение относительно интересующей Вас совокупности.

Пример 1

Испытание гипотезы по выборочному среднему – одна выборка

Университет внедрил новую кураторскую систему, которая нацелена на то, чтобы сократить число студентов, отчисляемых с курсов. Основываясь на опыте прошлых лет, администратор полагает, что в среднем будет 12 отчислений с курса без такой системы. Через год после внедрения системы была взята выборка 40 курсов, среднее число отчислений составило 11.175 со стандартным отклонением 2.308. Администратор желает знать, действительно ли количество отчислений существенно уменьшилось после внедрения кураторской системы.

№	количество отчисленных с курса
1	10
2	8
3	14
4	11
5	14
6	8
7	11
8	13
9	11
10	7
11	13
12	10
13	11
14	14
15	9
16	15
17	11
18	9
19	10
20	9
21	14
22	10
23	17
24	13
25	15
26	9
27	13
28	12
29	10

Среднее число отчислений на курс – 12, то есть H_0: $\mu=12$

H_1: $\mu<12$

В этом примере испытание одностороннее, потому что администратор интересуется, уменьшилось ли количество отчислений. Имеются следующие статистики выборки:

$\bar{x}=11.175$, $s=2.308$, $n=40$

Уровень значимости α примем 5%.

Проверочная статистика

$$t=\frac{(\bar{x}-\mu)}{\frac{s}{\sqrt{n}}}=\frac{(11.175-12)}{\frac{2.308}{\sqrt{40}}}=-2{,}2607$$

Число степеней свободы $n-1=40-1=39$

Критическое значение проверочной статистики находится по таблице Стьюдента

Уровень значимости нужно найти в верхней строке – заголовок столбца, а число степеней свободы в левом столбце – название строки. Наше значение находится на пересечении столбца 0.05 и строки 40 (наиболее близкое значение к 39). Коэффициент равен 1.6839.

α / **ν**	…	**0.05**	**0.025**	…
1	…	6.3138	12.7062	…
…	…	…	…	…
35	…	1.6896	2.0301	…
40	…	$\boxed{1{,}6839}$	2.0211	…
45	…	1.6794	2.0141	…

Так как абсолютное значение проверочной статистики больше критического значения,

$$|-2{,}2607|>|1.6839|$$

то нулевая гипотеза H_0 отвергается.

Вы можете утверждать, что, основываясь на выборочных данных, при уровне значимости 5% есть основания полагать, что среднее значение отчисленных студентов на один курс меньше 12.

Пример.2

Испытание гипотезы по выборочным средним – независимые выборки
Двухвыборочный t-тест с одинаковыми дисперсиями

Лектор по курсу бизнес-статистики пытается определить, действительно ли те студенты, кто ранее получили по математике уровень А (свыше 80 баллов), набирают по его предмету больше баллов, чем остальные студенты, которые имели уровень ниже А. Две выборки были отобраны случайно и получены следующие результаты

группа с уровнем А

№	балл
1	44
2	61
3	70
4	46
5	57
6	52
7	31
8	67
9	65
10	73
11	79
12	66
13	37
14	62
15	57
16	77
17	62
18	52
19	64
20	55
21	53
22	34
23	58
24	78
25	63
26	49
27	49
28	47

Среди имеющих уровень А по математике:

$$n_1 = 40 \quad \bar{x}_1 = \frac{\sum x}{n_1} = 56{,}500 \qquad s_1 = \sqrt{\frac{\sum (x-\bar{x}_1)^2}{n_1 - 1}} = 14{,}313$$

Среди не имеющих уровень А по математике:

$$n_2 = 75 \quad \bar{x}_2 = \frac{\sum x}{n_2} = 51{,}320 \qquad s_2 = \sqrt{\frac{\sum (x-\bar{x}_2)^2}{n_2 - 1}} = 14{,}943$$

Замечание. В задачах этого типа – с двумя независимыми выборками – нам придется предположить, что стандартные отклонения генеральных совокупностей σ_1 и σ_2 примерно равны.

В нашем примере стандартные отклонения выборок s_1=14,313 и s_2 = 14,943 достаточно близки и мы предполагаем $\sigma_1 = \sigma_2$.

Гипотезы:

$$H_0: \mu_1 = \mu_2$$

Среднее количество баллов, полученных в группе студентов с уровнем математики А существенно не отличается от среднего количества баллов, полученных в группе студентов с уровнем математики ниже А.

$$H_1: \mu_1 > \mu_2$$

В группе студентов с уровнем математики А среднее количество баллов больше, чем в группе студентов с уровнем математики ниже А..

Проведем испытание на уровне значимости 5%.

Проверочная статистика

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{56.500 - 51.320}{\sqrt{\frac{216{,}932}{40} + \frac{216{,}932}{75}}} = 1{,}7963$$

где $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)} =$

$$\frac{(40-1)\cdot 14.313^2 + (75-1)\cdot 14{,}943^2}{(40+75-2)} = 216{,}932$$

число степеней свободы

$$n_1 + n_2 - 2 = 40 + 75 - 2 = 113$$

остальные

№	балл
1	23
2	58
3	88
4	31
5	56
6	54
7	75
8	68
9	68
10	42
11	41
12	42
13	41
14	66
15	80
16	64
17	53
18	53
19	48
20	61
21	55
22	33
23	23
24	63
25	37
26	56

№	балл
38	25
39	76
40	52
41	48
42	55
43	27
44	57
45	54
46	26
47	58
48	38
49	71
50	67
51	49
52	47
53	66
54	41
55	50
56	56
57	53
58	45
59	71
60	29
61	52
62	46
63	50

Критическое значение для проверочной статистики берется из таблицы Стьюдента. На уровне значимости 5% для одностороннего испытания оно равно 1.6577.

ν \ α	…	0.05	0.025	…
…	…	…	…	…
100	…	1.6602	1.9840	…
120	…	$\boxed{1.6577}$	1.9799	…
…	…	…	…	…

Значение проверочной статистики больше, чем критическое значение.

$$|1{,}7963| > |1.6577|$$

Поэтому, нулевая гипотеза отклоняется.

Таким образом, на уровне значимости 5% имеются убедительные свидетельства того, что среднее количество баллов, получаемых по бизнес-статистике студентами с уровнем математики А существенно отличается от среднего количества баллов, получаемых студентами с уровнем математики ниже А.

Пример 3

Испытание гипотезы по парным данным – зависимые выборки

Парный двухвыборочный t-тест

Менеджер измеряет некоторый характерный показатель работы персонала до и после изменения системы мотивации. Тридцать проверенных сотрудников показали следующие результаты.

№	до	после	Различие x_d
1	38	35	3
2	33	33	0
3	31	31	0
4	29	29	0
5	32	30	2
6	32	32	0
7	39	32	7
8	39	33	6
9	31	31	0
10	29	27	2
11	35	35	0
12	29	26	3
13	42	38	4
14	32	36	-4
15	27	29	-2
16	36	34	2
17	32	32	0
18	35	35	0
19	37	35	2
20	31	30	1

Гипотезы:

H_0: $\mu_d = 0$, где μ_d – предполагаемое среднее различие.

Нулевая гипотеза состоит в том, что нет разницы в показателе до и после изменения системы.

H_1: $\mu_d \neq 0$

Альтернативная гипотеза полагает, что показатель изменился после внедрения системы, не уточняя в какую сторону.

Испытание двустороннее. Проведем его на уровне значимости 5%. Проверочная статистика

$$t = \frac{(\overline{x}_d - \mu_d)}{\frac{s_d}{\sqrt{n}}},$$

где x_d – различия, $\overline{x}_d$ – среднее различие в парах значений, s_d – стандартное отклонение различий.

Для этого примера

$\overline{x}_d$ =1.27 $\quad$ s_d =2.45 $\quad$ μ_d =0

$$t = \frac{(1{,}27 - 0)}{\frac{2{,}45}{\sqrt{30}}} = 2.8337$$

Критическое значение берется из таблицы Стьюдента, оно равно 2,0452

ν \ α	…	0.05	0.025	…
…	…	…	…	…
28	…	1.7011	2.0484	…
29	…	1.6991	$\boxed{2.0452}$	…
…	…	…	…	…

Значение проверочной статистики больше, чем критическое значение,

$$|2{,}8337| > |2{,}0452|$$

поэтому нулевая гипотеза отклоняется.

Менеджер может утверждать, что на уровне значимости 5% есть основания полагать, что показатель изменился.

Среднее значение показателя «до» $\overline{x}_1$=32.8, «после» $\overline{x}_2$=31.53, т.е. показатель уменьшился.

Обратите внимание. Менеджер мог бы попробовать использовать в этом примере **Двухвыборочный t-тест с одинаковыми дисперсиями**, считая выборки независимыми. Но при том же уровне значимости не смог бы отвергнуть нулевую гипотезу и доказать, что изменения произошли. **Парный двухвыборочный t-тест** является более сильным.

EXCEL КОМАНДЫ

EXCEL предлагает встроенную функцию=СТЬЮДРАСПОБР(α;число степеней свободы). Эта функция дает критическое значение из распределения Стьюдента для двустороннего испытания.

Для **примера 7.1** функция =СТЬЮДРАСПОБР(0,1;39)= 1,6849.

В примере 7.1 было использовано значение из таблицы 1,6839 для 40 степеней свободы.

Обратите внимание, если вы проводите одностороннее испытание, то уровень значимости в функции СТЬЮДРАСПОБР нужно удвоить.

В меню Сервис – Анализ Данных есть команды для проверки гипотез по двум выборкам

Для **примера 7.2** следует выбрать в меню Сервис – Анализ Данных - Двухвыборочный t-тест с одинаковыми дисперсиями

Его результат будет выведен как таблица

	Переменная 1	*Переменная 2*
Среднее	56,5	51,32
Дисперсия	204,8717949	223,3016216
Наблюдения	40	75
Объединенная дисперсия	216,940885	
Гипотетическая разность средних	0	
df	113	
t-статистика	1,79626647	
P(T<=t) одностороннее	0,037562031	
t критическое одностороннее	1,658449946	
P(T<=t) двухстороннее	0,075124063	
t критическое двухстороннее	1,981179594	

было получено s_p^2 =216,932

число степеней свободы

было получено t=1,7963

критическое значение 1.6577 ранее было взято для числа степеней свободы 115

Делать вывод, сравнивая проверочную t-статистику с критическим значением, все равно предстоит пользователю.

Для **примера .3** следует выбрать в меню Сервис – Анализ Данных - Двухвыборочный одинаковыми дисперсиями

Его результат будет выведен как таблица

	Переменная 1	*Переменная 2*
Среднее	32,8	31,53333
Дисперсия	15,2	11,22299
Наблюдения	30	30
Корреляция Пирсона	0,782008	
Гипотетическая разность средних	0	
df	29	
t-статистика	2,833439	
P(T<=t) одностороннее	0,004148	
t критическое одностороннее	1,699127	
P(T<=t) двухстороннее	0,008295	
t критическое двухстороннее	2,045231	

было получено t=2,8337

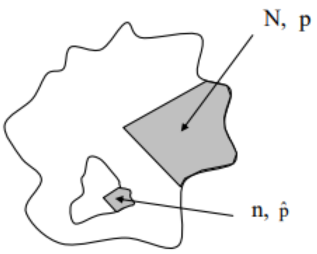
критическое значение из таблицы 2,0452

ИСПЫТАНИЕ ГИПОТЕЗЫ ПО ВЫБОРОЧНОЙ ДОЛЕ

Существуют две типичные ситуации

Испытание гипотезы по выборочной доле – одна выборка

Известно, какова доля особых элементов в генеральной совокупности. Взяв выборку, нужно проверить, не изменилась ли она.



Проверочная статистика $z = \dfrac{(\hat{p} - p)}{\sqrt{\dfrac{p(1-p)}{n}}}$

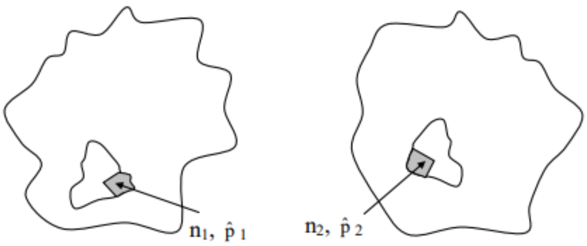
Условие: если выборка маленькая (n<30), то должно выполняться np≥5 и n(1-p) ≥5

Примеры задач

- По результатам продолжительного наблюдения за технологическим процессом известно, что доля несоответствующей продукции составляет 4%. Проведена реорганизация процесса. Взята выборка, и определен % бракованных изделий. Можно ли полагать, что доля брака уменьшилась?
- По результатам многолетних наблюдений доля сотрудников, которые берут больничный лист более 5 раз в году, составляет 1%. Введена новая система профилактики заболеваний. За первый год действия новой системы показатель уменьшился до 0,9%. Есть ли основания полагать, что доля часто болеющих сотрудников снизилась?

Испытание гипотезы по выборочным долям – две выборки

Имеются две выборки, в каждой из которых есть доля особых элементов. Нужно проверить, существует ли различие между долями особых элементов в совокупностях, из которых выборки взяты.



Проверочная статистика

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}},$$

где $\bar{p}$ – доля из объединенных выборок

Условие: обе выборки большие (n_1 >30 и n_2>30).

Примеры задач

- Внутренние аудиторы компании проводят в подразделении выборочную проверку документации. Некоторое количество документов признано оформленными неправильно. Аудиторы дали рекомендации по улучшению ведения документации. Следующая плановая проверка снова выявила некоторое количество неправильных документов. Можно ли полагать, что доля несоответствующих документов уменьшилась?
- Служба безопасности проводит проверку соблюдения правил внутреннего распорядка. В подразделении А численностью 500 человек выявлено 8 опоздавших, в подразделении В численностью 1000 человек – 12 опоздавших. Есть ли основания утверждать, что уровень нарушений в подразделении А выше?

Пример .4

Испытание гипотезы по выборочной доле – одна выборка

В поликлиниках города предлагается новая прививка от гриппа. Чтобы оценить ее эффективность, впоследствии ведется тщательный учет всех заболевших. Среди тех, кто не делал прививку (основная часть населения) заболевшие составляют 12,2 человек на 100 жителей города. Среди тех, кто сделал прививку, заболевшие составляют 7,6 человек из 100. Всего в городе сделано 512 прививок. Достаточно ли доказательств того, что люди, делающие прививку болеют гриппом реже?

Нулевая гипотеза: доля заболевших среди тех, кто делает прививки, такая же, как и среди тех, кто прививок не делает.

$$H_0: p=0.122$$

Альтернативная гипотеза: доля заболевших среди тех, кто делает прививки, меньше, чем среди тех, кто прививок не делает

$$H_1: p<0,122$$

Испытание одностороннее, проведем его на 5%-ном уровне значимости.

Проверочная статистика

$$z = \frac{(\hat{p} - p)}{\sqrt{\frac{p(1-p)}{n}}} = \frac{(0,076 - 0,122)}{\sqrt{\frac{0,122(1-0,122)}{512}}} = -3,180$$

Критическое значение для проверочной статистики находим из таблицы нормального распределения. Для уровня значимости 5% (0,05) при одностороннем испытании критическое значение приблизительно 1.645

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07214	.07078	.06944	.06811
1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938

Поскольку абсолютное значение проверочной статистики больше критического значения

$$|-3,180| < |1,645|,$$

то нулевая гипотеза отвергается. Существуют убедительные доказательства, что среди тех, кто делает прививку, доля болеющих ниже.

Пример .5

Испытание гипотезы по выборочным долям – две выборки

Внутренние аудиторы большой компании интересуются системой обработки счетов. Они взяли случайную выборку объемом $n_1 = 50$ счетов и проверяют их подробно. Четыре из них оказались неправильно оформленными. Тогда аудиторы предложили некоторые модификации в процедуре. Дав определенное время для внедрения новых процедур, аудиторы затем провели случайную выборку объемом $n_2 = 60$ счетов. Теперь они обнаружили три неправильных счета. Имеется ли какое-либо основание предполагать, что новые процедуры уменьшают ошибки?

Нулевая гипотеза предполагает, что две выборки случайно взяты из двух совокупностей с равными долями дефектов:

$$H_0: p_1 = p_2 = p$$

Альтернативная гипотеза предполагает, что новые процедуры сократили долю ошибок.

$$H_1: p_1>p_2,$$

Испытание одностороннее, проведем его на 5%-ном уровне значимости.

Доля неправильных документов в выборках

$$\hat{p}_1 = \frac{4}{50} = 0.08 \quad \text{и} \qquad \hat{p}_2 = \frac{3}{60} = 0.05$$

Оценка доли дефектных документов в генеральной совокупности достигается комбинированием долей двух выборок.

$$\bar{p} = \frac{4+3}{50+60} = 0.0636$$

Проверочная статистика

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.08 - 0.05)}{\sqrt{0.0636(1-0.0636)\left(\frac{1}{50} + \frac{1}{60}\right)}} = 0{,}642$$

Критическое значение такое же, как и в предыдущем примере 1,645.

Поскольку абсолютное значение проверочной статистики меньше критического значения

$$|0{,}642| < |1{,}645|,$$

то результат не существенен на 5%-ном уровне. Факты согласуются с H_0 на данном уровне значимости. У нас нет причины предполагать, что модификации в системе обработки документов сократили ошибки.

Обратите внимание. Если аудиторы увеличат выборку, то, возможно, им удастся опровергнуть нулевую гипотезу.

EXCEL КОМАНДЫ

EXCEL предлагает встроенную функцию =НОРМСТОБР(1-α). Эта функция дает критическое значение из нормального распределения для одностороннего испытания.

Для **примеров .4 и .5** функция = НОРМСТОБР (0,95)= 1,64485.

Обратите внимание, если бы на уровне значимости 5% проводилось двустороннее испытание, то критическое значение

= НОРМСТОБР (0,975)= 1,959963≅1.96

Контрольные вопросы.

1. Что такое нулевая гипотеза в тестах для численных переменных?
2. Как формулируется альтернативная гипотеза для теста Манна-Уитни
3. В чем различие в применении t-тестов и ранговых тестов. Приведите примеры ранговых тестов
4. Где можно найти проверочную статистику