Практическое занятие 11: Статистические гипотезы

Цель занятия:

- изучить типы ошибок;
- освоить этапы проверки статистических гипотез.

Понятие статистической гипотезы

Пусть исследуется некоторый признак **статистической совокупности**. Успеваемость студентов, продолжительность жизни, качество бензина. Всё, что можно «оцифровать» и посчитать.

Как проводится исследование?

Обычно так: из **генеральной совокупности** извлекается **репрезентативная выборка**, и на основании изучения этой выборки делается вывод обо всей совокупности.

Это основной метод математической статистики и называется он **выборочным методом**. В зависимости от исследования, могут проводиться неоднократные выборки, выборки из нескольких ген. совокупностей, да и вообще анализироваться произвольные статистические данные.

И в результате анализа этих данных появляются мысли, которые оформляются в **статистические гипотезы**.

Статистической называют гипотезу о законе распределения статистической совокупности **либо** о числовых параметрах известных (!) распределений.

Например:

- рост танкистов распределен **нормально**;
- **дисперсии** стрельбы двух танковых дивизий равны между собой, при этом известно, что точность стрельбы распределена нормально;
- из многочисленных ранее проведённых исследований.

В первом случае выдвигается гипотеза о законе распределения, во втором – о числовых характеристиках двух распределений, закон которых известен.

Откуда взялись эти гипотезы?

В первом случае была проведена **выборка** танкистов (например, 100 человек) и в результате её исследования появилось обоснованное предположение, что рост ВСЕХ танкистов распределён нормально.

Во втором случае исследовались **выборочные** данные по точности стрельбы двух дивизий, в результате чего возник интерес проверить – а одинакова ли генеральная результативность, или же какая-то дивизия стреляет точнее?

В обеих гипотезах речь идёт о генеральных совокупностях, и выдвигаются эти гипотезы на основании анализа **выборочных** данных. Это распространенная схема, но она не единственна, бывают и другие статистические гипотезы.

Выдвигаемую гипотезу называют **нулевой** и обозначают через H_0. Обычно это наиболее очевидная и правдоподобная гипотеза (хотя это вовсе не обязательно). И в противовес к ней рассматривают **альтернативную** или **конкурирующую** гипотезу H_1.

В рассмотренных выше примерах альтернативные гипотезы очевидны (отрицают нулевую), но существуют и другие варианты, так, например, к гипотезе H_0 :**генеральная средняя** нормально распределённой совокупности равна $a=10$, можно сформулировать разные конкурирующие гипотезы H_1: $a\neq 10$, H_1: $a>10$, H_1: $a<10$, H_1: $a=11$.

Поскольку нулевая гипотеза выдвигается на основании выборочных данных, то она может оказаться как правильной, так и неправильной. И поэтому она подлежит **статистической проверке**.

Проверка осуществляется с помощью **статистических критериев** – это специальные случайные величины, которые принимают различные действительные значения.

В результате проверки нулевая гипотеза либо принимается, либо отвергается в пользу альтернативной. При этом есть риск допустить ошибки двух типов:

Ошибка первого рода - ***отвергнута правильная*** гипотеза H_0. Вероятность допустить такую ошибку называют **уровнем значимости** и обозначают буквой α («альфа»).

Ошибка второго рода – ***принята неправильная*** гипотеза H_0. Вероятность совершить эту ошибку обозначают буквой β («бета»).

Значение $1-\beta$ называют **мощностью критерия** – это ***вероятность отвержения неправильной гипотезы***.

Уровень значимости задаётся исследователем самостоятельно, наиболее часто выбирают значения $\alpha=0{,}1$, $\alpha=0{,}05$, $\alpha=0{,}01$.

И тут возникает мысль, что чем меньше «альфа», тем вроде бы лучше. Но это только вроде: **при уменьшении вероятности** α - отвергнуть правильную гипотезу **растёт вероятность** β - принять неверную гипотезу (при прочих равных условиях).

Поэтому перед исследователем стоит задача грамотно подобрать соотношение вероятностей α и β, при этом учитывается **тяжесть последствий**, которые повлекут за собой та и другая ошибки.

Пример: пользователь **X** создал п/я. По умолчанию, H_0 – он считается добропорядочным пользователем. Так считает антиспам фильтр. И вот **X** отправляет письмо.

После чего фильтр может совершить ошибку двух типов:

1) ошибочно отклонить гипотезу H_0 (счесть нормальное письмо за спам и **X** за спаммера)

2) ошибочно принять гипотезу H_0 (хотя **X** - спаммер).

Какая ошибка более «тяжелая»? Письмо от X может быть ОЧЕНЬ важным для адресата, и поэтому при настройке фильтра целесообразно уменьшить уровень значимости α, пожертвовав вероятностью β, в результате чего в основной ящик будут чаще попадать письма со спамом.

Существует примеры, где наоборот – более тяжкие последствия влечёт ошибка 2-го рода, и вероятность α следует увеличить (в пользу уменьшения вероятности β).

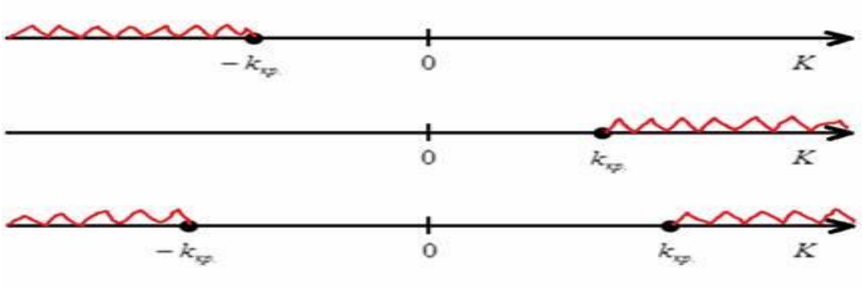
Этапы процесса проверки статистической гипотезы:

1) Обработка выборочных данных и выдвижение основной H_0 и конкурирующей H_1 гипотез.

2) Выбор статистического критерия K. Это **непрерывная случайная величина**, принимающая различные действительные значения.

3) Выбор уровня значимости α.

4) Нахождение **критического значения** $k_{кр}$ – это значение случайной величины K, которое зависит от выбранного уровня значимости α и опционально от других параметров. Критическое значение определяет **критическую область**. Она бывает левосторонней, правосторонней и двусторонней (красная штриховка):



Критическая область – это ***область отвержения нулевой гипотезы***. Незаштрихованную область называют ***областью принятия гипотезы***.

5) Далее на основании выборочных данных рассчитывается ***наблюдаемое значение критерия***: $k_{набл}$.

6) Решение:

- если $k_{набл}$ в критическую область НЕ попадает, то гипотеза H_0 на уровне значимости α принимается. Таким образом с вероятностью α рисковали отвергнуть правильную гипотезу. Однако не нужно думать, что нулевая гипотеза доказана и 100% правильна, ведь существует вероятность β – того совершена ошибка 2-го рода (принята неверная гипотеза).
- если $k_{набл}$ попадает в критическую область, то гипотеза H_0 на уровне значимости α отвергается *(при этом, если, например,* α=0,05*, то в среднем в 5 случаев из 100 мы отвергнем правильную гипотезу, т.е. совершим ошибку 1-го рода).*

Гипотеза о генеральной средней нормального распределения

Постановка задачи: предполагается, что генеральная средняя $\boldsymbol{a}$ нормального распределения равна некоторому значению $\boldsymbol{a_0}$.

Это нулевая гипотеза: $\boldsymbol{H_0}\quad \boldsymbol{a = a_0}$.

Для проверки гипотезы на уровне значимости $\boldsymbol{\alpha}$ проводится **выборка** объема $\boldsymbol{n}$ и рассчитывается **выборочная средняя** $\overline{x}_в$.

Исходя из полученного значения и специфики той или иной задачи, можно сформулировать следующие конкурирующие гипотезы:

1) $\boldsymbol{H_1 : a < a_0}$
2) $\boldsymbol{H_1 : a > a_0}$
3) $\boldsymbol{H_1 : a \neq a_0}$
4) $\boldsymbol{H_1 : a = a_1}$, где $\boldsymbol{a_1}$ – конкретное альтернативное значение *генеральной средней*.

При этом возможны две принципиально разные ситуации:

а) генеральная дисперсия $\boldsymbol{\sigma^2}$ известна.

Тогда в качестве *статистического критерия* $\boldsymbol{K}$ рассматривают случайную величину $U = \dfrac{(\overline{X} - a_0)\sqrt{n}}{\sigma}$, где $\overline{X}$ – случайное значение *выборочной средней*. Почему случайное? Потому что в разных выборках мы будем получать разные значения $\overline{x}_в$, и заранее предугадать это значение невозможно.

5) Далее находим *критическую область*.
   - для конкурирующих гипотез $\boldsymbol{H_1 : a < a_0}$ и $\boldsymbol{H_1 : a = a_1}$ (случай $\boldsymbol{a_1 < a_0}$) строится *левосторонняя область*;
   - для гипотез $\boldsymbol{H_1 : a > a_0}$ и $\boldsymbol{H_1 : a = a_1}$ (случай $\boldsymbol{a_1 > a_0}$)) – *правосторонняя*;
   - для гипотезы $\boldsymbol{H_1 : a \neq a_0}$ – *двусторонняя* – т. к. конкурирующее значение генеральной средней может оказаться как больше, так и меньше $\boldsymbol{a_0}$-го.

Чтобы найти критическую область нужно отыскать критическое значение $\boldsymbol{u_{кр}}$.

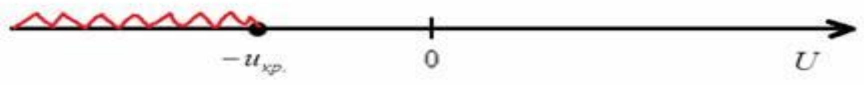
Оно определяется из соотношения

- $\boldsymbol{\Phi(u_{кр}) = (1-2\alpha)/2}$ – для односторонней области *(лево- или право-);*
- $\boldsymbol{\Phi(u_{кр}) = (1-\alpha)/2}$ – для двусторонней области, где $\boldsymbol{\alpha}$ – выбранный уровень значимости, а $\boldsymbol{\Phi(u)}$ –функция Лапласа.

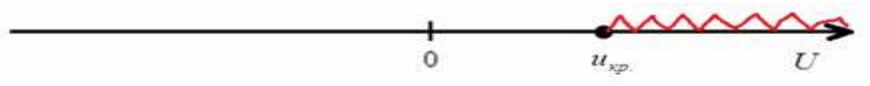
Теперь на основании выборочных данных рассчитываем *наблюдаемое значение критерия*: $U = \dfrac{(\overline{x}_в - a_0)\sqrt{n}}{\sigma}$.

Результаты:

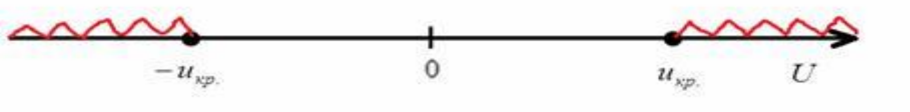
1) Для левосторонней критической области.
Если $\boldsymbol{u_{набл}}$>-$\boldsymbol{u_{кр}}$, то гипотеза $\boldsymbol{H_0}$ на уровне значимости $\boldsymbol{\alpha}$ принимается.
Если $\boldsymbol{u_{набл}}$<-$\boldsymbol{u_{кр}}$, то гипотеза $\boldsymbol{H_0}$ на уровне значимости $\boldsymbol{\alpha}$ отвергается.



2) Для правосторонней критической области.
Если $\boldsymbol{u_{набл}}$<$\boldsymbol{u_{кр}}$, то гипотеза $\boldsymbol{H_0}$ на уровне значимости $\boldsymbol{\alpha}$ принимается.
Если $\boldsymbol{u_{набл}}$>$\boldsymbol{u_{кр}}$, *(красный цвет)* – гипотеза $\boldsymbol{H_0}$ отвергается.



3) Двусторонняя критическая область.
Если -$\boldsymbol{u_{кр}}$<$\boldsymbol{u_{набл}}$<$\boldsymbol{u_{кр}}$ *(незаштрихованный интервал)*, то гипотеза $\boldsymbol{H_0}$ принимается, в противном случае – отвергается.



Условие принятия гипотезы часто записывают компактно – с помощью модуля: |$\boldsymbol{u_{набл}}$|<$\boldsymbol{u_{кр}}$.

б) генеральная дисперсия $\boldsymbol{\sigma^2}$ НЕ известна.

В этом случае остаётся ориентироваться на **исправленную выборочную дисперсию $\boldsymbol{S^2}$** и критерий $T = \dfrac{(\overline{X} - a_0)\sqrt{n}}{S}$, где $\overline{X}$ – случайное значение *выборочной средней* и $\boldsymbol{S}$ – соответствующее **исправленное стандартное отклонение**.

Задание 1

Из нормальной генеральной совокупности с известной дисперсией σ^2=3,2 извлечена выборка объёма $\boldsymbol{n}$=25 и по ней найдена выборочная средняя $\overline{x}_в$=19,3. Требуется на уровне значимости 0,01 проверить нулевую гипотезу $\boldsymbol{H_0}$: $\boldsymbol{a}$=20 против конкурирующей гипотезы $\boldsymbol{H_1}$: $\boldsymbol{a}$=19.

Методика выполнения

Прежде чем приступить к решению, пару слов о смысле такой задачи. Есть **генеральная совокупность** с известной **дисперсией** и есть веские основания полагать, что **генеральная средняя** равна 20 (нулевая гипотеза). В результате выборочной проверки получена **выборочная средняя** 19,3, и возникает **вопрос**: это результат случайный или же генеральная средняя и на самом деле меньше 20? – в частности, равна 19 (конкурирующая гипотеза).

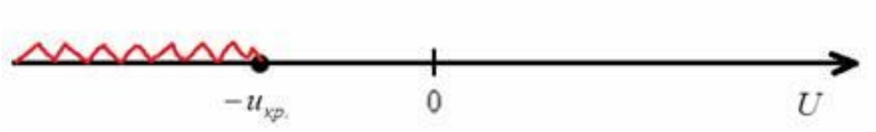
Решение: по условию, известна **генеральная дисперсия** σ^2=3,2, поэтому для проверки гипотезы $\boldsymbol{H_0}$: $\boldsymbol{a=a_0}$=20 используем случайную величину $U = \dfrac{(\overline{X}-a_0)\sqrt{n}}{\sigma}$.

Найдём критическую область. Для этого нужно найти критическое значение. Так как конкурирующее значение $\boldsymbol{H_1}$: $\boldsymbol{a=a_1}$=19 **меньше** чем $\boldsymbol{a_0}$=20, то критическая область будет ***левосторонней***.

Критическое значение для уровня значимости $\boldsymbol{\alpha}$=0,01 определим из соотношения:

$\boldsymbol{\Phi(u_{кр}) = (1-2\alpha)/2}$ =(1-2*0,01)/2=0,49

По таблице значений функции Лапласа или с помощью MS Excel определяем, что этому значению функции соответствует аргумент $\boldsymbol{u_{кр}} \approx 2{,}33$. Таким образом, при $\boldsymbol{u<-u_{кр}}$ *(красная критическая область)* нулевая гипотеза отвергается, а при $\boldsymbol{u>-u_{кр}}$ – принимается:



В данном случае $\boldsymbol{-u_{кр}} \approx -2{,}33$.

Вычислим наблюдаемое значение критерия:

$$u_{набл} = \frac{(\overline{X}-a_0)\sqrt{n}}{\sigma} = \frac{(19{,}3-20)\sqrt{25}}{\sqrt{3{,}2}} \approx -1{,}96$$

$\boldsymbol{u_{набл}>-u_{кр}}$, поэтому **на уровне значимости** $\boldsymbol{\alpha}$=0,01 **нулевую гипотезу** $\boldsymbol{H_0}$: $\boldsymbol{a}$=20 **принимаем**.

Такой, вроде бы неожиданный результат, объясняется тем, что генеральное стандартное отклонение достаточно велико $\sigma = \sqrt{3,2} \approx 1,79$, а поэтому нет оснований отвергать «главное» значение a_0=20 (несмотря на то, что выборочная средняя $\overline{x}_в$=19,3 гораздо ближе к конкурирующему значению a_1=19). Иными словами, такое значение выборочной средней, вероятнее всего, объясняется естественным разбросом **вариант x_i.**

Ответ: на уровне значимости 0,01 нулевую гипотезу принимаем.

Что означает *на уровне значимости 0,01»*? Это означает, что мы с 1%-ной вероятностью рисковали отвергнуть нулевую гипотезу, при условии, что она действительно справедлива. Однако не нужно забывать, что на самом деле она может быть и неверной и существует β -вероятность того, мы приняли неправильную гипотезу.

Варианты задания 1

$H_0 : a = 20$		$H_1 : a = 19$			$H_0 : a = 20$		$H_1 : a = 19$		
№ вар.	**Дисперсия**	***n***	$\overline{x}_в$	***Ур. зн.***	***№ вар.***	**Диспе**	***n***	$\overline{x}_в$	***Ур.***
1	**3,2**	**29**	**19,2**	**0,01**	**14**	**3,1**	**21**	**19,0**	**0,05**
2	**3,1**	**27**	**19,0**	**0,01**	**15**	**3,1**	**26**	**18,9**	**0,05**
3	**3,0**	**23**	**18,9**	**0,01**	**16**	**3,3**	**22**	**19,2**	**0,05**
4	**3,2**	**21**	**19,2**	**0,01**	**17**	**3,2**	**28**	**19,2**	**0,05**
5	**3,2**	**20**	**18,9**	**0,01**	**18**	**3,2**	**28**	**19,0**	**0,05**
6	**3,1**	**21**	**18,8**	**0,01**	**19**	**3,0**	**28**	**19,3**	**0,05**
7	**2,9**	**30**	**19,3**	**0,01**	**20**	**3,1**	**27**	**19,2**	**0,05**
8	**3,2**	**29**	**19,0**	**0,01**	**21**	**3,3**	**28**	**19,1**	**0,05**
9	**3,1**	**27**	**18,8**	**0,01**	**22**	**3,2**	**29**	**18,9**	**0,05**
10	**3,0**	**23**	**19,0**	**0,01**	**23**	**3,1**	**21**	**19,3**	**0,05**
11	**3,3**	**23**	**18,9**	**0,01**	**24**	**3,2**	**24**	**18,9**	**0,05**
12	**2,9**	**21**	**19,0**	**0,01**	**25**	**2,9**	**27**	**19,0**	**0,05**
13	**3,0**	**23**	**19,2**	**0,05**	**26**	**3,1**	**22**	**19,1**	**0,05**

Задание 2

По результатам $\boldsymbol{n}$=5 измерений температуры в печи найдено $\overline{x}_в$ =256^{0}С. Предполагается, что ошибка измерения есть нормальная случайная величина с $\boldsymbol{\sigma}$=6^{0}С. Проверить на уровне значимости $\boldsymbol{\alpha}$=0,05 гипотезу $\boldsymbol{H_0}$: $\boldsymbol{a}$=250 ^{0}С против конкурирующей гипотезы $\boldsymbol{H_1}$: $\boldsymbol{a}$>250 ^{0}С.

Методика выполнения

Сначала разберём, в чём жизненность этой ситуации. Есть печка. Для нормального технологического процесса нужна температура 250 градусов. Для проверки этой нормы 5 раз измерили температуру, получили 256 градусов. Из многократных предыдущих опытов известно, что среднеквадратическая погрешность измерений составляет 6 градусов (она обусловлена погрешностью самого термометра, случайными обстоятельствами проверки и т.д.)

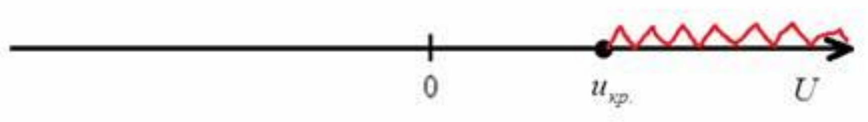
И здесь не понятно, почему выборочный результат (256 градусов) получился больше нормы – то ли температура действительно выше и печь нуждается в регулировке, то ли это просто погрешность измерений, которую можно не принимать во внимание.

Решение: по условию, известно **ген. среднее квадратическое отклонение** $\boldsymbol{\sigma}$=6^{0}С, поэтому для проверки гипотезы $\boldsymbol{H_0}$: $\boldsymbol{a}$=250 ^{0}С используем случайную величину $U = \frac{(\overline{X} - a_0)\sqrt{n}}{\sigma}$.

Найдём критическую область. Так как в конкурирующей гипотезе $\boldsymbol{H_1}$: $\boldsymbol{a}$>250 ^{0}С речь идёт о больших значениях температуры, то эта область будет ***правосторонней***.

Критическое значение для уровня значимости $\boldsymbol{\alpha}$=0,05 определим из соотношения $\boldsymbol{\Phi(u_{кр}) = (1\text{-}2\alpha)/2}$ = (1-2*0,05)/2= 0,45.

По таблице значений функции Лапласа определяем, что $\boldsymbol{u_{кр}}$ ≈1,645. Таким образом, при $\boldsymbol{u>u_{кр}}$ (критическая область) нулевая гипотеза отвергается, а при $\boldsymbol{u<u_{кр}}$ – принимается:



Вычислим наблюдаемое значение критерия:

$$u_{набл} = \frac{(\overline{x}_в - a_0)\sqrt{n}}{\sigma} = \frac{(256-250)\sqrt{5}}{6} \approx 2{,}236.$$

$u_{набл} > u_{кр}$, поэтому **на уровне значимости** α=0,05 **нулевую гипотезу** H_0: a=250^{0}С **отвергаем**.

Как бы сказали статистики, выборочный результат $\overline{x}_в$=256^{0}С ***статистически значимо*** отличается от нормативного значения 250^{0}С, и печь нуждается в регулировке (для уменьшения температуры).

Ответ: **на уровне значимости** α=0,05 **нулевую гипотезу** H_0: a=250^{0}С **отвергаем**.

Ещё раз осмыслим – что означает *«на уровне значимости 0,05»*? Это означает, что с вероятностью 5% мы отвергли правильную гипотезу (совершили ошибку 1-го рода).

И тут остаётся взвесить риск – насколько критично чуть-чуть уменьшить температуру (если мы всё-таки ошиблись и температура на самом деле в норме). Если даже небольшое уменьшение температуры недопустимо, то имеет смысл провести повторное, более качественное исследование: увеличить количество замеров n, использовать более совершенный термометр, улучшить условия эксперимента и т.д.

Варианты задания 2

$H_0: a$=250 ^{0}С H_1: a>250 ^{0}С

№ вар.	n	СКО	$\overline{x}_в$	Ур. зн.
1	6	5	255	0,10
2	4	4	247	0,10
3	9	7	248	0,10
4	7	7	256	0,10
5	10	7	251	0,10
6	4	6	253	0,10
7	8	6	247	0,10
8	5	4	252	0,10
9	9	4	251	0,10
10	9	7	254	0,10
11	10	7	254	0,10
12	5	5	256	0,10
13	10	7	247	0,05

$H_0: a$=250 ^{0}С H_1: a>250 ^{0}С

№ вар.	n	СКО	$\overline{x}_в$	Ур. зн.
14	5	7	252	0,05
15	10	5	247	0,05
16	4	4	258	0,05
17	10	4	248	0,05
18	7	5	249	0,05
19	8	5	251	0,05
20	8	4	249	0,05
21	6	7	247	0,05
22	8	7	251	0,05
23	6	5	247	0,05
24	6	4	250	0,05
25	4	4	248	0,05
26	8	6	256	0,05

Задание 3

Средний вес таблетки сильнодействующего лекарства (номинал) должен быть равен 0,5 мг. Выборочная проверка n=100 выпущенных таблеток показала, что средний вес таблетки равен $\overline{x}_в$=0,508 мг. Многократными предварительными опытами по взвешиванию таблеток, изготавливаемых фармацевтическим заводом, установлено, что вес таблеток распределен нормально со средним квадратическим отклонением σ=0,11 мг. Требуется на уровне значимости α=0,1 проверить гипотезу о том, что средний вес таблеток действительно равен H_0: a=0,5 мг**.**

Методика выполнения

Следует рассмотреть как конкурирующую гипотезу H_1: a>0,5, так и гипотезу H_1: $a\neq$=0,5. И в самом деле – ведь полученное значение $\overline{x}_в$ =0,508 мг является случайным и в другой выборке оно может запросто оказаться и меньше чем 0,5.

Решение: поскольку известно ген. стандартное отклонение, то для проверки гипотезы H_0: a=0,5 используем случайную величину $U = \frac{(\overline{X} - a_0)\sqrt{n}}{\sigma}$.

а) Рассмотрим конкурирующую гипотезу H_1: a>0,5. Так как альтернативные значения генеральной средней больше чем 0,5, то находим ***правостороннюю*** критическую область.

Критическое значение для уровня значимости α=0,1 определим из соотношения $\boldsymbol{\Phi(u_{кр}) = (1\text{-}2\alpha)/2}$ = (1-2*0,1)/2= 0,4

При $u<u_{кр}$ гипотеза принимается, при $u>u_{кр}$ (критическая область) – отвергается.

Вычислим наблюдаемое значение критерия:

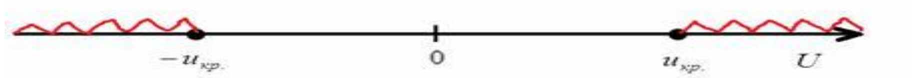
$$u_{набл} = \frac{(\overline{x}_в - a_0)\sqrt{n}}{\sigma} = \frac{(0{,}508 - 0{,}5)\sqrt{100}}{0{,}11} \approx 0{,}73$$

$u_{набл}>u_{кр}$, таким образом, на уровне значимости 0,1 гипотезу H_1: a>0,5 отвергаем, а гипотезу H_0: a=0,5 принимаем.

б) Рассмотрим конкурирующую гипотезу H_1: $a\neq$0,5. В данном случае критическая область ***двусторонняя***.

Критическое значение для уровня значимости α=0,1 определим из соотношения $\boldsymbol{\Phi(u_{кр}) = (1\text{-}\alpha)/2}$ = (1-0,1)/2= 0,45 $\Rightarrow u_{кр} \approx$ 1,64.

При $-u_{кр}<u<u_{кр}$ гипотеза принимается, а при $u>|u_{кр}|$ (красная критическая область) – отвергается:



Наблюдаемое значение критерия $u_{набл}\approx$0,73 вычислено в предыдущем пункте, и оно попадает в область принятия гипотезы.

Ответ: в обоих случаях нулевую гипотезу на уровне значимости 0,1 принимаем.

Кстати, это тот самый пример, где ошибка 2-го рода *(ошибочное принятие неверной нулевой гипотезы)*, может повлечь гораздо более тяжелые последствия (опасную передозировку). Поэтому в такой ситуации лучше включить паранойю и увеличить уровень значимости α – при этом мы будем чаще отвергать правильную нулевую гипотезу (совершать ошибку 1-го рода), но зато перестрахуемся и проведём более тщательное исследование.

Можно ли **одновременно** уменьшить вероятности ошибок 1-го и 2-го рода (α и β)? Да можно. Если увеличить объём выборки. Что вполне логично.

Алгоритм решения полностью сохраняется:

Варианты задания 3

$H_0 : a$ =0,5 мг					$H_0 : a$ =0,5 мг				
№ вар.	***n***	***СКО***	$\bar{x}_в$	***Ур. зн.***	**№ вар.**	***n***	***СКО***	$\bar{x}_в$	***Ур. зн.***
1	**115**	**0,10**	**0,512**	**0,10**	**14**	**86**	**0,12**	**0,507**	**0,05**
2	**111**	**0,11**	**0,514**	**0,10**	**15**	**108**	**0,13**	**0,499**	**0,05**
3	**88**	**0,12**	**0,514**	**0,10**	**16**	**114**	**0,11**	**0,504**	**0,05**
4	**106**	**0,11**	**0,507**	**0,10**	**17**	**114**	**0,12**	**0,503**	**0,05**
5	**116**	**0,15**	**0,511**	**0,10**	**18**	**85**	**0,14**	**0,51**	**0,05**
6	**115**	**0,10**	**0,505**	**0,10**	**19**	**104**	**0,14**	**0,514**	**0,05**
7	**116**	**0,11**	**0,508**	**0,10**	**20**	**97**	**0,11**	**0,507**	**0,05**
8	**101**	**0,15**	**0,5**	**0,10**	**21**	**120**	**0,13**	**0,499**	**0,05**
9	**95**	**0,13**	**0,506**	**0,10**	**22**	**89**	**0,11**	**0,505**	**0,05**
10	**116**	**0,10**	**0,497**	**0,10**	**23**	**92**	**0,14**	**0,51**	**0,05**
11	**97**	**0,11**	**0,515**	**0,10**	**24**	**83**	**0,15**	**0,506**	**0,05**
12	**120**	**0,11**	**0,511**	**0,10**	**25**	**87**	**0,14**	**0,51**	**0,05**
13	**92**	**0,11**	**0,495**	**0,05**	**26**	**107**	**0,12**	**0,503**	**0,05**

Задание 4

На основании n=7 измерений найдено, что средняя высота сальниковой камеры равна $\bar{x}_в$=51 мм. Многократными предварительными опытами установлено, что высота сальниковой камеры распределена нормально со средним квадратическим отклонением σ=1,2 мм. В предположении о нормальном распределении проверить на уровне значимости α=0,05 гипотезу H_0: a=50 мм против конкурирующей гипотезы H_1: $a \neq 50$ мм.

Методика выполнения

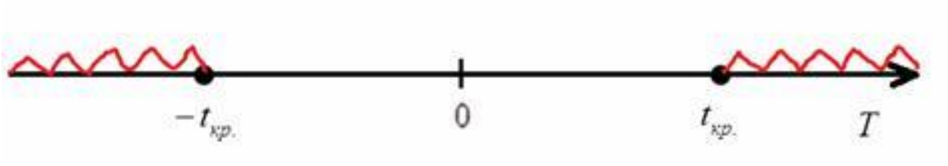
Смысл задачи: 7 раз измерили высоту этой камеры, получили среднее значение 51 мм и за неимением генеральной дисперсии вычислили **исправленную выборочную дисперсию**. Согласно норме, высота должна равняться 50 мм – эту гипотезу и проверяем.

Решение:

Конкурирующая гипотеза имеет вид ***H₁***: ***a***≠50 мм, а значит, речь идёт о ***двусторонней*** критической области.

Критическое значение для уровня значимости ***α***=0,05 определим из соотношения $\boldsymbol{\Phi(u_{кр}) = (1-\alpha)/2}$ = (1-0,05)/2= 0,475 ⇒ $\boldsymbol{u_{кр}} \approx$ 1,96.

При $\boldsymbol{-u_{кр}<u_{набл}<u_{кр}}$ нулевая гипотеза принимается, и вне этого интервала (в критической области при $t > |t_{кр}|$) – отвергается:



Вычислим наблюдаемое значение критерия:

$$u_{набл} = \frac{(\overline{x}_в - a_0)\sqrt{n}}{\sigma} = \frac{(51-50)\sqrt{7}}{1{,}2} \approx 2{,}2$$

Полученное значение **не попало** в область принятия гипотезы (-1,96<$\boldsymbol{u_{набл}}$<1,96), поэтому **на уровне значимости 0,05 нулевую гипотезу отвергаем**.

Ответ: на уровне значимости 0,05 гипотезу ***H₀***: ***a***=50 мм **отвергаем**.

Варианты задания 4

H_0: a=50 мм			H_1: a≠50 мм		H_0: a=50 мм			H_1: a≠50 мм	
№ вар.	n	СКО	$\overline{x}_в$	Ур. зн.	№ вар.	n	СКО	$\overline{x}_в$	Ур. зн.
1	8	1,00	50	0,10	14	5	1,05	50	0,05
2	7	1,27	51	0,10	15	6	1,13	52	0,05
3	8	1,21	48	0,10	16	9	1,03	52	0,05
4	10	1,15	48	0,10	17	9	1,14	49	0,05
5	9	1,10	52	0,10	18	5	1,06	52	0,05
6	10	1,06	50	0,10	19	8	1,27	51	0,05
7	10	1,03	52	0,10	20	5	1,16	52	0,05
8	8	1,15	48	0,10	21	6	1,14	48	0,05
9	9	1,02	49	0,10	22	8	1,03	51	0,05
10	9	1,08	52	0,10	23	9	1,10	48	0,05
11	7	1,12	48	0,10	24	10	1,12	48	0,05
12	9	1,05	51	0,10	25	8	1,09	52	0,05
13	10	1,21	52	0,05	26	10	1,26	48	0,05