МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Московский авиационный институт
(национальный исследовательский университет)»

Кафедра «Моделирование систем и информационные технологии»

Гипотезы о средних. Сравнение двух средних.

Методические указания к практическому занятию
по дисциплине "Математическая статистика"

Составители: Егорова Ю.Б.
Мамонов И.М.

МОСКВА 2020

ВВЕДЕНИЕ

Цель практического занятия – изучить способы проверки статистической гипотезы о сравнении двух выборочных средних (или двух математических ожиданий) нормального распределения для двух случаев (генеральная дисперсия известна; генеральная дисперсия неизвестна).

ПРИМЕР 1. Для проверки эффективности новой технологии отобраны две группы рабочих. В первой группе рабочих численностью 50 человек, где применялась новая технология, выборочная средняя выработка составила 85 изделий. Во второй группе численностью 70 человек выборочная средняя составила 78 изделий. Генеральные дисперсии в группах соответственно равны 100 и 74 (изделия)^2. При уровне значимости 0,05 необходимо выяснить значимо ли влияние новой технологии на среднюю выработку.

РЕШЕНИЕ. По условию n_1=50; $\bar{x}=85$ изд.; $D(X)=100$ изд.^{2}; n_2=70; $\bar{y}=78$ изд.; $D(Y)=74$ изд.^{2}, α=0,05. Выдвигаем нулевую гипотезу H_o: $M(X)=M(Y)$. Относительно альтернативной гипотезы возможны два случая: а) $M(X)\neq M(Y)$; б) H_1: $M(X)>M(Y)$ (так как $\bar{x}>\bar{y}$). Рассмотрим эти случаи.

а) H_o: $M(X)=M(Y)$

H_1: $M(X)\neq M(Y)$.

В этом случае строят двустороннюю критическую область.

<u>Порядок проверки нулевой гипотезы:</u>

1) по выборке определяем наблюдаемое значение критерия:

$$Z_{набл} = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{D(X)}{n_1}+\frac{D(Y)}{n_2}}} = \frac{85-78}{\sqrt{\frac{100}{50}+\frac{74}{70}}} = 4{,}00.$$

2) по таблице функции Лапласа (см. Приложение 1) определяем критическое значение критерия $z_{крит} = 1{,}96$ из равенства $\phi(z_{крит}) = \frac{1}{2} - \frac{\alpha}{2} = \frac{1}{2} - \frac{0{,}05}{2} = 0{,}475.$

3) Так как $|Z_{набл}| = 4{,}00 > z_{крит} = 1{,}96$, то нулевая гипотеза отвергается и принимается альтернативная.

б) H_o: $M(X)=M(Y)$.

H_1: $M(X)>M(Y)$.

В этом случае строят правостороннюю критическую область.

<u>Порядок проверки нулевой гипотезы:</u>

1) по выборке определяем наблюдаемое значение критерия $Z_{набл} = 4{,}00.$

2) по таблице функции Лапласа (см. Приложение 1) определяем критическое значение критерия $z_{крит}$ =1,645 из равенства $\phi(z_{крит}) = \frac{1}{2} - \alpha = \frac{1}{2} - 0{,}05 = 0{,}45.$

3) Так как $|Z_{набл}| = 4{,}00 > z_{крит} = 1{,}645$, то нулевая гипотеза отвергается и принимается альтернативная.

Таким образом, с вероятностью 0,95 можно утверждать, что выборочные средние различаются значимо и, следовательно, новая технология позволяет повысить среднюю выработку рабочих.

ПРИМЕР 2. Технологи механосборочного цеха считают, что применение нового резца позволит сократить время обработки детали. Пять деталей были изготовлены старым резцом. Среднее время обработки одной детали составило 3,3 мин с «исправленной» дисперсией 0,25 мин^2. Шесть деталей были изготовлены новым резцом. Среднее время обработки одной детали составило 2,48 мин с «исправленной» дисперсией 0,108 мин^2. При уровне значимости 0,05 проверьте, позволило ли использование нового типа резцов сократить время обработки детали.

РЕШЕНИЕ. По условию n_1=5; $\bar{x} = 3{,}3$ мин; $S_x^2 = 0{,}25$ мин.^{2}; n_2=6; $\bar{y} = 2{,}48$ мин; $S_y^2 = 0{,}108$ мин^2, α=0,05. Генеральные дисперсии неизвестны, но будем полагать, что они одинаковы*.

Выдвигаем нулевую гипотезу H_o: $M(X)=M(Y)$. Относительно альтернативной гипотезы возможны два случая: а) $M(X)\neq M(Y)$; б) H_1: $M(X)>M(Y)$ (так как $\bar{x} > \bar{y}$). Рассмотрим эти случаи.

а) H_o: $M(X)=M(Y)$

H_1: $M(X)\neq M(Y)$.

В этом случае строят двустороннюю критическую область.

<u>Порядок проверки нулевой гипотезы:</u>

1) по выборке определяем наблюдаемое значение критерия $T_{набл}$:

$$T_{набл} = \frac{\bar{x}-\bar{y}}{\sqrt{(n_1-1)S_x^2+(n_2-1)S_y^2}}\sqrt{\frac{n_1\cdot n_2\cdot(n_1+n_2-2)}{n_1+n_2}} =$$

$$= \frac{3{,}3-2{,}48}{\sqrt{(5-1)\cdot 0{,}25+(6-1)\cdot 0{,}108}}\sqrt{\frac{5\cdot 6\cdot(5+6-2)}{5+6}} = 3{,}27.$$

2) по таблице критических точек распределения Стьюдента (см. Приложение 2) определяем критическое значение критерия $t_{крит}$=2,26 для односторонней критической области в зависимости от уровня значимости α=0,05 и числа степеней свободы $k=n_1+n_2$-2=5+6-2=9.

3) Так как $T_{набл} = 3{,}26 > t_{крит} = 2{,}26$, то нулевая гипотеза отвергается и принимается альтернативная.

б) H_o: $M(X)=M(Y)$.

H_1: $M(X)>M(Y)$.

В этом случае строят правостороннюю критическую область.

<u>Порядок проверки нулевой гипотезы:</u>

1) по выборке определяем наблюдаемое значение критерия $T_{набл} = 3{,}26$.

2) по таблице критических точек распределения Стьюдента (см. Приложение 2) определяем критическое значение критерия $t_{крит}$=1,83 для односторонней критической области в зависимости от уровня значимости α=0,05 и числа степеней свободы $k=n_1+n_2$-2=5+6-2=9.

3) Так как $T_{набл} = 3{,}26 > t_{крит} = 1{,}83$, то нулевая гипотеза отвергается и принимается альтернативная.

Таким образом, с вероятностью 0,95 можно утверждать, что выборочные средние различаются значимо и, следовательно, использование нового типа резцов позволило сократить время обработки детали.

* Предварительно необходимо проверить гипотезу о равенстве генеральных дисперсий.

ПРИМЕР 3. Годовой оборот 4 бирж в регионе *A* составил 120 тыс. у.е., а в регионе *B* годовой оборот 5 бирж – 125 тыс. у.е. Исправленная выборочная дисперсия в регионе *A* оказалась равной 30 тыс. у.е., в регионе *B* – 20 тыс. у.е. Можно ли при уровне значимости 0,05 утверждать, что средний оборот бирж в регионе *A* меньше, чем в регионе *B*?

РЕШЕНИЕ. По условию n_1=4; $\bar{x} = 120$ тыс. у.е.; $S_x^2 = 30$ тыс. у.е.^{2}; n_2=5; $\bar{y} = 125$ тыс. у.е.; $S_y^2 = 20$ тыс. у.е.^{2}. Генеральные дисперсии неизвестны, но будем полагать, что они одинаковы*.

Выдвигаем нулевую гипотезу H_o: $M(X)=M(Y)$.

Выдвигаем альтернативную гипотезу H_1: $M(X)<M(Y)$.

В этом случае строят левостороннюю критическую область.

<u>Порядок проверки нулевой гипотезы:</u>

1) по выборке определяем наблюдаемое значение критерия $T_{набл}$:

$$T_{набл} = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)S_x^2 + (n_2 - 1)S_y^2}} \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}} =$$

$$= \frac{12 - 12{,}5}{\sqrt{(4-1)\cdot 3 + (5-1)\cdot 2}} \sqrt{\frac{4 \cdot 5 \cdot (4+5-2)}{4+5}} = -47{,}8.$$

2) по таблице критических точек распределения Стьюдента (см. Приложение 2) определяем критическое значение критерия $t_{крит}$=1,89 для односторонней критической области в зависимости от уровня значимости α=0,05 и числа степеней свободы $k=n_1+n_2$-2=4+5-2=7.

* Предварительно необходимо проверить гипотезу о равенстве генеральных дисперсий.

3) Так как $T_{набл} = -47{,}8 < -t_{крит} = -1{,}89$, то нулевая гипотеза отвергается и принимается альтернативная.

Таким образом, с вероятностью 0,95 можно утверждать, что выборочные средние различаются значимо и, следовательно, средний оборот бирж в регионе *A* меньше, чем в регионе *B*.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. По двум независимым выборкам n_1=40 и n_2=50, извлеченным из нормальных генеральных совокупностей X и Y, найдены выборочные средние 130 и 140 соответственно. Генеральные дисперсии известны D(X)=80 и D(Y)=100. При уровне значимости 0,95 проверить нулевую гипотезу о равенстве математических ожиданий двух генеральных совокупностей.
2. По двум независимым выборкам n_1=12 и n_2=18, извлеченным из нормальных генеральных совокупностей X и Y, найдены выборочные средние 31,2 и 29,2 соответственно. Генеральные дисперсии неизвестны, но предполагается, что равны между собой. Определены исправленные дисперсии S_x^2=0,40 и S_y^2=0,84. При уровне значимости 0,95 проверить нулевую гипотезу о равенстве математических ожиданий двух генеральных совокупностей.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Сформулируйте основные принципы проверки гипотез.
2. Как проверяется гипотеза о равенстве двух математических ожиданий, если дисперсии известны?
3. Как проверяется гипотеза о равенстве двух математических ожиданий, если дисперсии неизвестны?

Приложение 1

Таблица значений функции Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-x^2/2} dx$

x	*Ф(x)*	*x*	*Ф(x)*	*x*	*Ф(x)*	*x*	*Ф(x)*
0,00	0,0000	0,42	0,1628	0,84	0,2995	1,26	0,3969
0,01	0,0040	0,43	0,1664	0,85	0,3023	1,27	0,3980
0,02	0,0080	0,44	0,1700	0,86	0,3051	1,28	0,3997
0,03	0,0120	0,45	0,1736	0,87	0,3078	1,29	0,4015
0,04	0,0160	0,46	0,1772	0,88	0,3106	1,30	0,4032
0,05	0,0199	0,47	0,1808	0,89	0,3133	1,31	0,4049
0,06	0,0239	0,48	0,1844	0,90	0,3159	1,32	0,4066
0,07	0,0279	0,49	0,1879	0,91	0,3186	1,33	0,4082
0,08	0,0319	0,50	0,1915	0,92	0,3212	1,34	0,4099
0,09	0,0359	0,51	0,1950	0,93	0,3238	1,35	0,4115
0,10	0,0398	0,52	0,1985	0,94	0,3264	1,36	0,4131
0,11	0,0438	0,53	0,2019	0,95	0,3289	1,37	0,4147
0,12	0,0478	0,54	0,2054	0,96	0,3315	1,38	0,4162
0,13	0,0517	0,55	0,2088	0,97	0,3340	1,39	0,4177
0,14	0,0557	0,56	0,2123	0,98	0,3365	1,40	0,4192
0,15	0,0596	0,57	0,2157	0,99	0,3389	1,41	0,4207
0,16	0,0636	0,58	0,2190	1,00	0,3413	1,42	0,4222
0,17	0,0675	0,59	0,2224	1,01	0,3438	1,43	0,4236
0,18	0,0714	0,60	0,2257	1,02	0,3461	1,44	0,4251
0,19	0,0753	0,61	0,2291	1,03	0,3485	1,45	0,4265
0,20	0,0793	0,62	0,2324	1,04	0,3508	1,46	0,4279
0,21	0,0832	0,63	0,2357	1,05	0,3531	1,47	0,4292
0,22	0,0871	0,64	0,2389	1,06	0,3554	1,48	0,4306
0,23	0,0910	0,65	0,2422	1,07	0,3577	1,49	0,4319
0,24	0,0948	0,66	0,2454	1,08	0,3599	1,50	0,4332
0,25	0,0987	0,67	0,2486	1,09	0,3621	1,51	0,4345
0,26	0,1026	0,68	0,2517	1,10	0,3643	1,52	0,4357
0,27	0,1064	0,69	0,2549	1,11	0,3665	1,53	0,4370
0,28	0,1103	0,70	0,2580	1,12	0,3686	1,54	0,4382
0,29	0,1141	0,71	0,2611	1,13	0,3708	1,55	0,4394
0,30	0,1179	0,72	0,2642	1,14	0,3729	1,56	0,4406
0,31	0,1217	0,73	0,2673	1,15	0,3749	1,57	0,4418
0,32	0,1255	0,74	0,2703	1,16	0,3770	1,58	0,4429
0,33	0,1293	0,75	0,2734	1,17	0,3790	1,59	0,4441
0,34	0,1331	0,76	0,2764	1,18	0,3810	1,60	0,4452
0,35	0,1368	0,77	0,2794	1,19	0,3830	1,61	0,4463
0,36	0,1406	0,78	0,2823	1,20	0,3849	1,62	0,4474
0,37	0,1443	0,79	0,2852	1,21	0,3869	1,63	0,4484
0,38	0,1480	0,80	0,2881	1,22	0,3883	1,64	0,4495
0,39	0,1517	0,81	0,2910	1,23	0,3907	1,65	0,4505
0,40	0,1554	0,82	0,2939	1,24	0,3925	1,66	0,4515
0,41	0,1591	0,83	0,2967	1,25	0,3944	1,67	0,4525

x	*Φ(x)*	*x*	*Φ(x)*	*x*	*Φ(x)*	*x*	*Φ(x)*
1,68	0,4535	1,91	0,4719	2,28	0,4887	2,74	0,4969
1,69	0,4545	1,92	0,4726	2,30	0,4893	2,76	0,4971
1,70	0,4554	1,93	0,4732	2,32	0,4898	2,78	0,4973
1,71	0,4564	1,94	0,4738	2,34	0,4904	2,80	0,4974
1,72	0,4573	1,95	0,4744	2,36	0,4909	2,82	0,4976
1,73	0,4582	1,96	0,4750	2,38	0,4913	2,84	0,4977
1,74	0,4591	1,97	0,4756	2,40	0,4918	2,86	0,4979
1,75	0,4599	1,98	0,4761	2,42	0,4922	2,88	0,4980
1,76	0,4608	1,99	0,4767	2,44	0,4927	2,90	0,4981
1,77	0,4616	2,00	0,4772	2,46	0,4931	2,92	0,4982
1,78	0,4625	2,02	0,4783	2,48	0,4934	2,94	0,4984
1,79	0,4633	2,04	0,4793	2,50	0,4938	2,96	0,4985
1,80	0,4641	2,06	0,4803	2,52	0,4941	2,98	0,4986
1,81	0,4649	2,08	0,4812	2,54	0,4945	3,00	0,49865
1,82	0,4656	2,10	0,4821	2,56	0,4948	3,20	0,49931
1,83	0,4664	2,12	0,4830	2,58	0,4951	3,40	0,49966
1,84	0,4671	2,14	0,4838	2,60	0,4953	3,60	0,499841
1,85	0,4678	2,16	0,4846	2,62	0,4956	3,80	0,499928
1,86	0,4686	2,18	0,4854	2,64	0,4959	4,00	0,499968
1,87	0,4693	2,20	0,4861	2,66	0,4961	4,50	0,499997
1,88	0,4699	2,22	0,4868	2,68	0,4963	5,00	0,499997
1,89	0,4706	2,24	0,4875	2,70	0,4965		
1,90	0,4713	2,26	0,4881	2,72	0,4967		

Приложение 2

Критические точки распределения Стьюдента

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)	
	0,1	0,05
1	6,31	12,70
2	2,92	4,30
3	2,35	3,18
4	2,13	2,78
5	2,01	2,57
6	1,94	2,45
7	1,89	2,36
8	1,86	2,31
9	1,83	2,26
10	1,81	2,23
11	1,80	2,20
12	1,78	2,18
13	1,77	2,16
14	1,76	2,14
15	1,75	2,13
16	1,75	2,12
17	1,74	2,11
18	1,73	2,10
19	1,73	2,09
20	1,73	2,09
…	…	…
	0,05	0,025
	Уровень значимости α (односторонняя критическая область)	