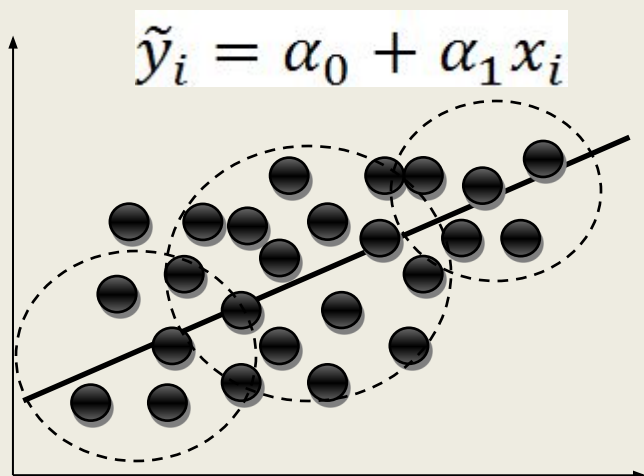


Оценка существенности уравнения регрессии и его параметров

1. Проверка статистических гипотез
2. Оценка существенности уравнения регрессии и его параметров

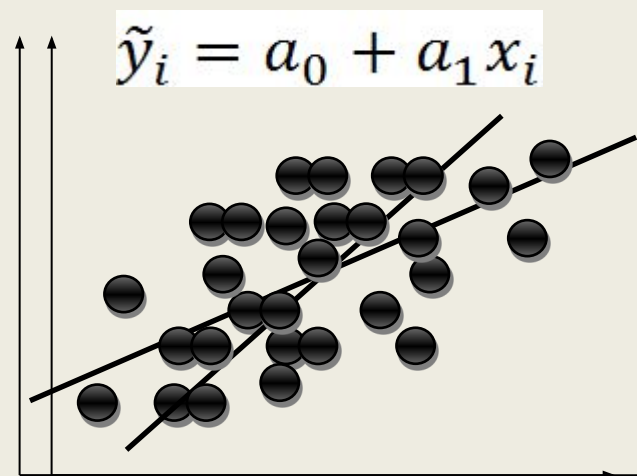
Проверка статистических гипотез

Суть проверки статистических гипотез



$$\tilde{y}_i = \alpha_0 + \alpha_1 x_i$$

Генеральная совокупность

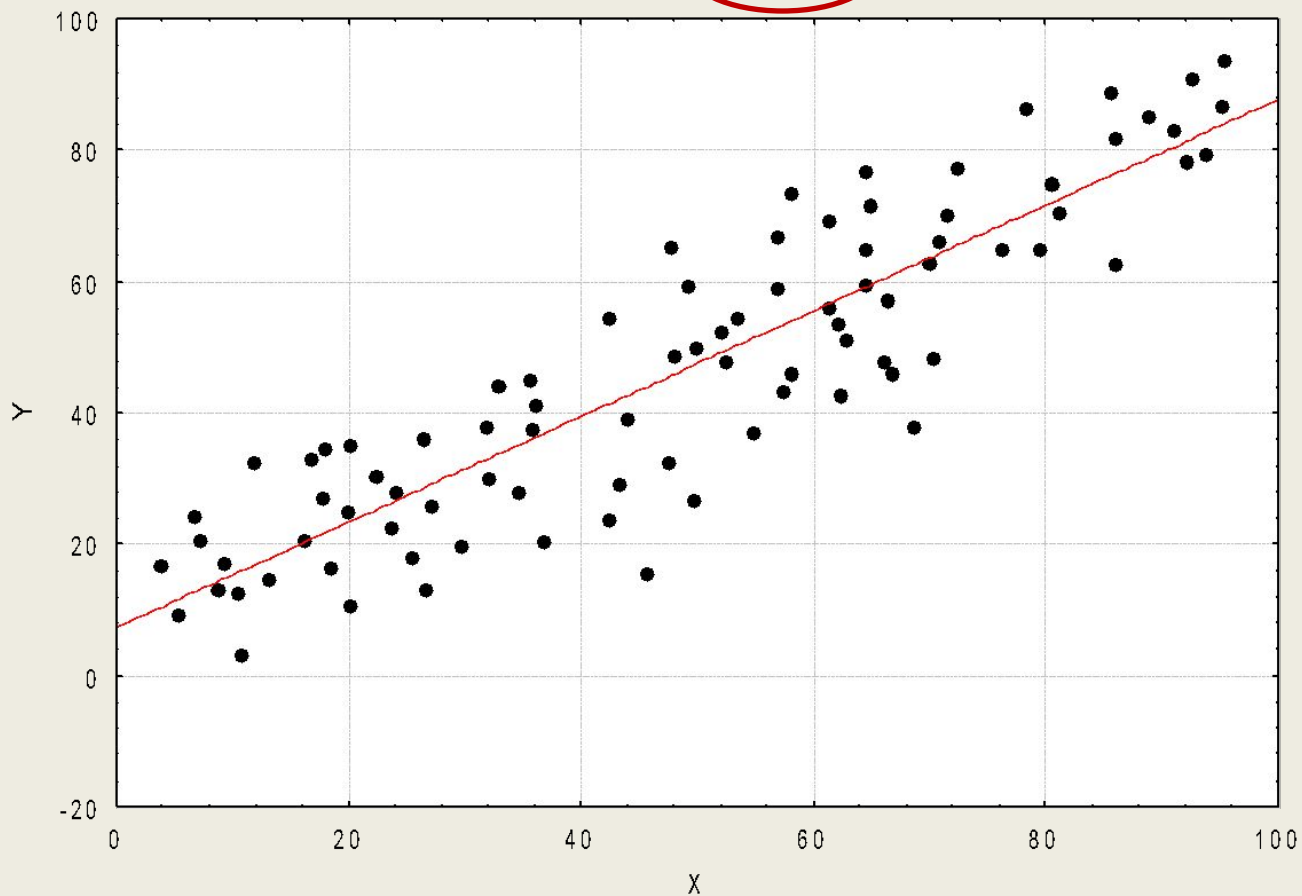


$$\tilde{y}_i = a_0 + a_1 x_i$$

Выборочная совокупность

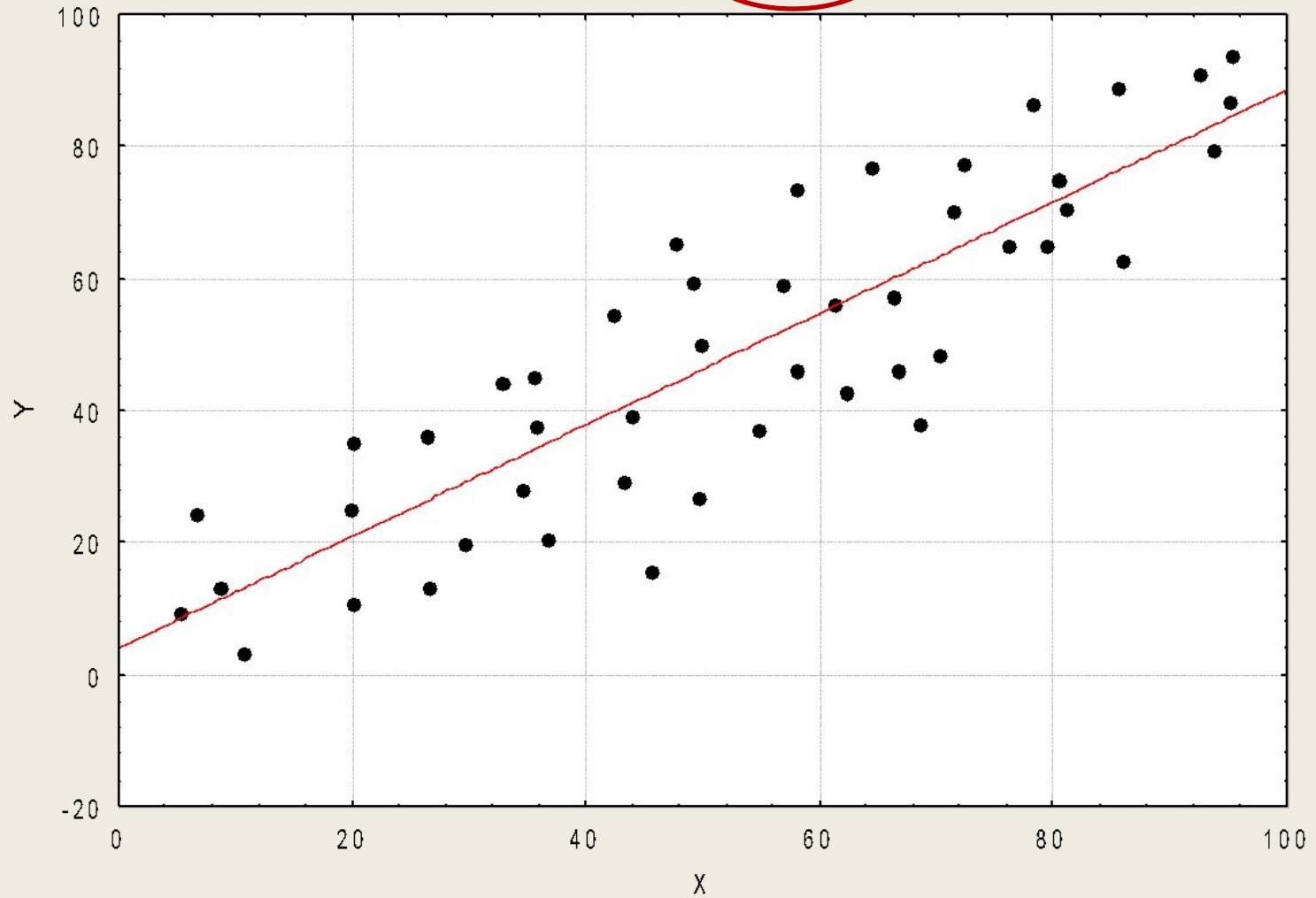
Генеральная совокупность

$$Y = 7,2897 + 0,8045x$$



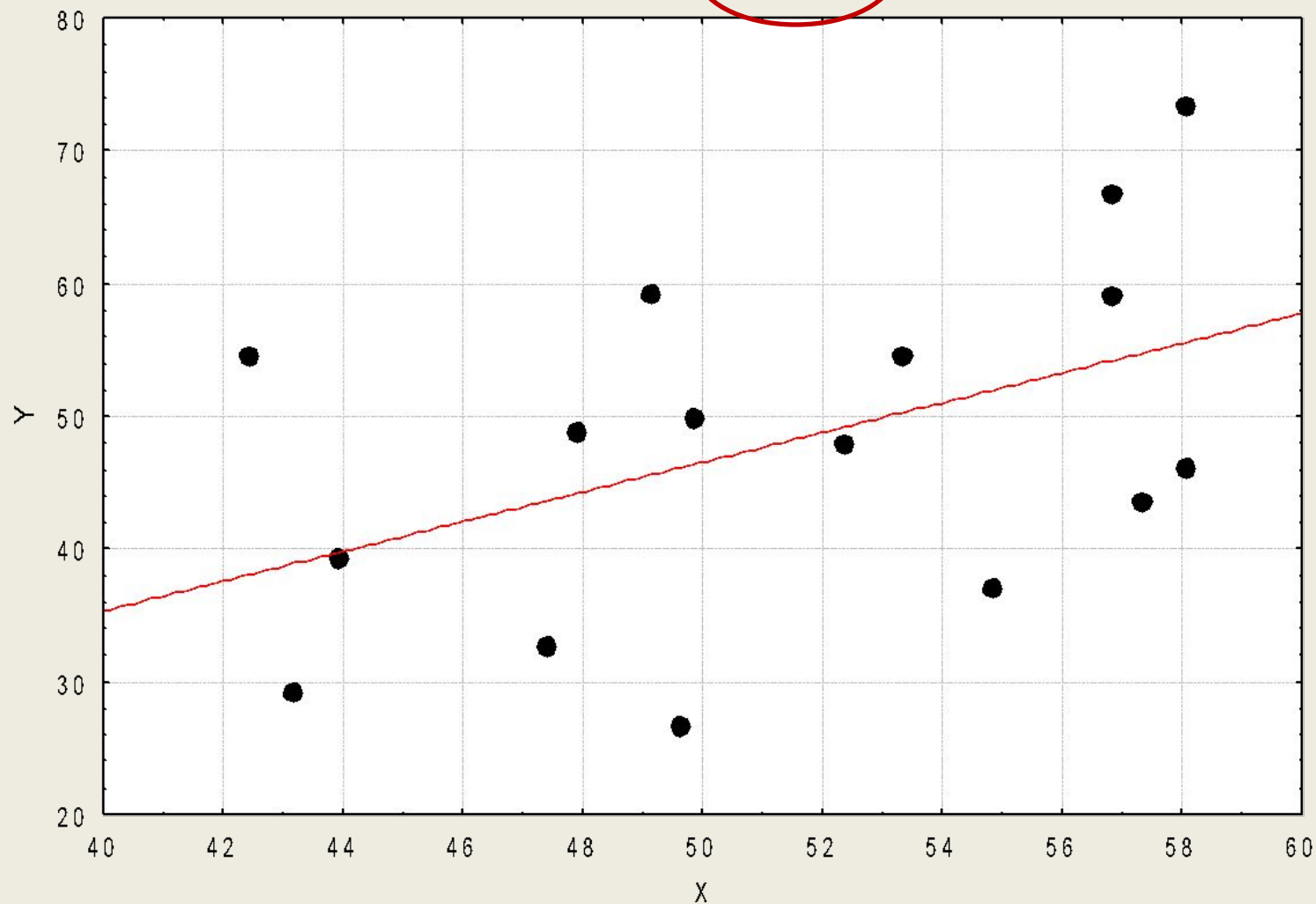
Выборочная совокупность

$$Y = 4,0254 + 0,8452 * X$$



Выборочная совокупность (диапазон X от **40** до **60**)

$$Y = -9,482 + 1,1205 * X$$



H_0 - нулевая гипотеза

H_1 – альтернативная гипотеза

Ошибка первого рода произойдет, когда будет принято решение отклонить гипотезу H_0 , хотя в действительности она оказывается верной.

Ошибка второго рода произойдет, когда будет принято решение не отклонять гипотезу H_0 хотя в действительности она будет неверна.

Результат проверки	Возможные состояния проверяемой гипотезы	
	Верна гипотеза H_0	Верна гипотеза H_1
Гипотеза H_0	Ошибка первого рода	Правильно, решение
Гипотеза H_0 не отклоняется	Правильное решение	Ошибка второго рода

Уровнем значимости называется вероятность ошибочного отклонения нулевой гипотезы. Или иными словами уровень значимости это вероятность ошибки первого рода при принятии решения.

Стандартные уровни статистической значимости

- ✓ низшим уровнем статистической значимости является уровень $P = 0,05$;
- ✓ достаточным - уровень $P = 0,01$;
- ✓ высшим - уровень $P = 0,001$.

Y - Доходы консолидированных бюджетов субъектов российской федерации , млн. руб.

X1 - Численность экономически активного населения, тыс. чел.

1% уровень значимости

	<i>Beta</i>	<i>Std.Err. of Beta</i>	<i>B</i>	<i>Std.Err. of B</i>	<i>t(12)</i>	<i>p-level</i>
<i>Intercept</i>			-11095,5	4649,618	-2,386	0,0344
<i>X1</i>	0,961	0,079	44,9	3,714	12,098	0,0000

отвергается
на 1% уровне
значимости

5% уровень значимости

	<i>Beta</i>	<i>Std.Err. of Beta</i>	<i>B</i>	<i>Std.Err. of B</i>	<i>t(12)</i>	<i>p-level</i>
<i>Intercept</i>			-11095,5	4649,618	-2,386	0,034
<i>X1</i>	0,961	0,079	44,9	3,714	12,098	0,000

отвергается
на 5% уровне
значимости

Правило принятия статистического вывода таково:

1 шаг. Выдвигается H_0 и H_1

2 шаг. Рассчитывают эмпирическую (фактическая или расчетная) статистику - $X_{\text{факт}}$

3 шаг. По таблице распределения находят критическое (табличное) значение статистики $X_{\text{кр}}$

4 шаг. Сравнивают $X_{\text{факт}}$ и $X_{\text{кр}}$

Оценка существенности уравнения регрессии и его параметров

Алгоритм построения t -критерия Стьюдента:

1 шаг.

$$H0: a_0 = 0; a_1 = 0; \dots a_k = 0$$

$$H1: a_0 \neq 0; a_1 \neq 0; \dots a_k \neq 0$$

2 шаг.

$$t_{a_0} = |a_0| \frac{\sqrt{n-2}}{\sigma_{\text{ост}}}$$

$$t_{a_1} = |a_1| \frac{\sqrt{n-2}}{\sigma_{\text{ост}}} \sigma_x$$

$$\sigma_{\text{ост}} = \sqrt{\sum (y_i - \tilde{y}_i)^2 / n}$$

$$\sigma_x = \sqrt{\sum (x_i - \bar{x})^2 / n}$$

3 шаг. =СТЬЮДРАСПОБР(вероятность; степени_свободы)

4 шаг.

$$t_{\text{факт.}} > t_{\text{табл.}}$$

Какой из параметров уравнения статистически значим?

Итоги регрессии для зависимой переменной: Y (Таблица данных1) R= ,88083329 R2= ,77586729 Скорректир. R2= ,75862631 F(1,13)=45,001 p<,00001 Станд. ошибка оценки: ,73939						
N=15	БЕТА	Стд.Ош. БЕТА	B	Стд.Ош. B	t(13)	p-уров.
Св.член			1,397	1,647	0,848	0,412
X1	0,881	0,131	0,041	0,006	6,708	0,000

Какой из параметров уравнения статистически значим?

Итоги регрессии для зависимой переменной: Y (Таблица данных1) R= ,98811246 R2= ,97636623 Скорректир. R2= ,97242727 F(2,12)=247,87 p<,00000 Станд. ошибка оценки: ,24990						
N=15	БЕТА	Стд.Ош. БЕТА	B	Стд.Ош. B	t(12)	p-уров.
Св.член			6,732	0,768	8,768	0,000
X1	0,193	0,081	0,009	0,004	2,370	0,035
X2	0,821	0,081	1,107	0,110	10,090	0,000

Критические значения t -критерия Стьюдента на уровне значимости 0,10; 0,05; 0,01 (двухсторонний)

Число средней свободы df	α		
	0,10	0,05	0,01
1	6,3138	12,7060	63,6570
2	2,9200	4,3027	9,9248
3	2,3534	3,1825	5,8409
4	2,1318	2,7764	4,6041
5	2,0150	2,5706	4,0321
6	1,9432	2,4469	3,7074
7	1,8946	2,3646	3,4995
8	1,8595	2,3060	3,3554
9	1,8331	2,2622	3,2498
10	1,8125	2,2281	3,1693
11	1,7959	2,2010	3,1058
12	1,7823	2,1788	3,0545
13	1,7709	2,1604	3,0123
...
∞	1,6449	1,9600	2,5758

Алгоритм построения F -критерий Фишера:

1 шаг. $H_0: a_1 = a_2 = \dots = a_k = 0$

шаг. $H_1: a_1 \neq a_2 \neq \dots \neq a_k \neq 0$

2 шаг.
$$F_{\text{факт.}} = \frac{\sum(\tilde{y}_i - \bar{y})^2 / m - 1}{\sum(y_i - \tilde{y}_i)^2 / n - m} = \frac{\sigma_R^2}{\sigma_E^2}$$

3 шаг. =FRASПОБР(вероятность; степени_свободы1;
степени_свободы2)

4 шаг.
$$F_{\text{факт.}} > F_{\text{табл.}}$$

Оценить статистическую значимость уравнения регрессии.

Дисперсионный анализ; ЗП: Y (Таблица данных1)					
Эффект	Сумма квадрат	сс	Средн. квадрат	F	p-уров.
Регресс.	30,960	2	15,480	247,874	0,000
Остатки	0,749	12	0,062		
Итого	31,709				

Таблица значений F -критерия Фишера на уровне значимости $\alpha=0,05$

v_2	v_1	1	2	3	4	5	...	∞
1		161,45	191,50	215,72	224,57	230,17	...	254,32
2		18,51	19,00	19,16	19,25	19,30	...	19,50
3		10,13	9,50	9,28	9,12	9,01	...	8,53
4		7,71	6,54	6,59	6,39	6,26	...	5,63
5		6,61	5,49	5,41	5,19	5,05	...	4,36
6		5,99	5,14	4,76	4,53	4,39	...	3,67
7		5,59	4,74	4,35	4,12	3,97	...	3,23
8		5,32	4,46	4,07	3,84	3,69	...	2,93
9		5,12	4,26	3,86	3,63	3,48	...	2,71
10		4,96	4,10	3,71	3,48	3,33	...	2,54
11		4,84	3,98	3,59	3,36	3,20	...	2,40
12			3,88	3,49	3,26	3,11	...	2,30
...		2,21
∞		3,84	2,99	2,60	2,37	2,21	...	1,00

**Лекция окончена
спасибо за внимание**

Следующая лекция
«Нелинейная регрессия»