

# Проверка значимости коэффициентов регрессии

Построено уравнение  $\hat{y} = \hat{a} \cdot x + \hat{b}$

Необходимо проверить значимость коэффициентов  $\hat{a}$  и  $\hat{b}$

Значимость коэффициента говорит о том, что он существенно отличается от нуля. Незначимость говорит о том, что можно считать, что коэффициент равен 0.

**Если коэффициент  $\hat{a}$  незначим, то переменная  $y$  не зависит от переменной  $x$ .**

Если коэффициент  $\hat{b}$  незначим, то прямая проходит через начало координат.

Более важным является проверка значимости коэффициента  $\hat{a}$ . Если он незначим, то построенное уравнение бессмысленно, им нельзя пользоваться для прогнозирования!

1. Рассчитываются стандартные ошибки коэффициентов регрессии по специальным формулам

$$S_a, S_b$$

Excel делает это автоматически при построении регрессии через Сервис-Анализ данных

2. Рассчитывается статистика критерия

$$T_a = \frac{\hat{a}}{s_a} \quad \text{для проверки значимости } a$$

$$T_b = \frac{\hat{b}}{s_b} \quad \text{для проверки значимости } b$$

Excel делает это автоматически при построении регрессии через  
Сервис-Анализ данных

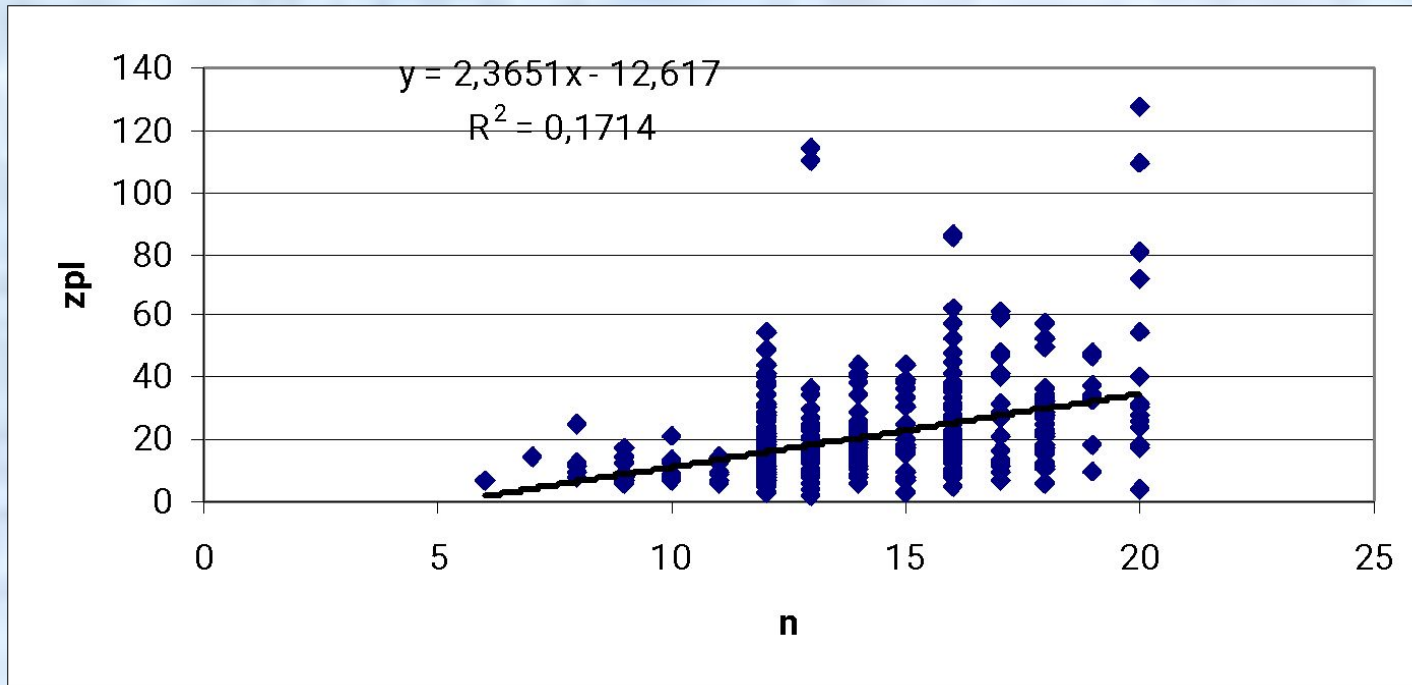
3. Рассчитываются Р-значения.

**Р-значение** - это вероятность того, что коэффициент не значим (можно считать, что он равен 0). При Р-значении больше 0,05 обычно считают, что коэффициент не значим, меньше 0,05 - значим.

Р-значения Excel рассчитывает автоматически.

# ПРИМЕР УРАВНЕНИЯ ПАРНОЙ РЕГРЕССИИ

$$Z_{pl} = -12,6188 + 2,3651N$$



Проверим значимость коэффициента перед N.

Незначимость этого коэффициента будет означать, что

уровень образования не влияет на заработную плату. Значимость - влияет



# ПРИМЕР УРАВНЕНИЯ ПАРНОЙ РЕГРЕССИИ

ВЫВОД ИТОГОВ					
<i>Регрессионная статистика</i>					
Множественный R	0,413988				
R-квадрат	0,171386				
Нормированный R	0,169846				
Стандартная ошибка	13,30384				
Наблюдения	540				
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>значимость</i>
Регрессия	1	19695,21	19695,21	111,2774	8,96E-24
Остаток	538	95221,75	176,9921		
Итого	539	114917			
<i>Коэффициент</i>					
<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Верхние 95%</i>	<i>Нижние 95%</i>	
Y-пересечение	-12,6168	3,118382	-4,04596	5,97E-05	-18,7425
n	2,365094	0,224205	10,54881	8,96E-24	1,92467

стандартные ошибки

t-статистики для гипотез

P-значения

Так как P-значения меньше 0,05, оба коэффициента значимы. Число лет обучения влияет на заработную плату.