Министерство науки и высшего образования Российской Федерации

**КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ ЭНЕРГЕТИЧЕСКИЙ УНИВЕРСИТЕТ**

**Отчет   
по лабораторной работе №2  
дисциплины «Статистика»  
Нелинейные модели парной регрессии**

Выполнил: Соловьёв Л.А.

Группа: ПИ-1-22

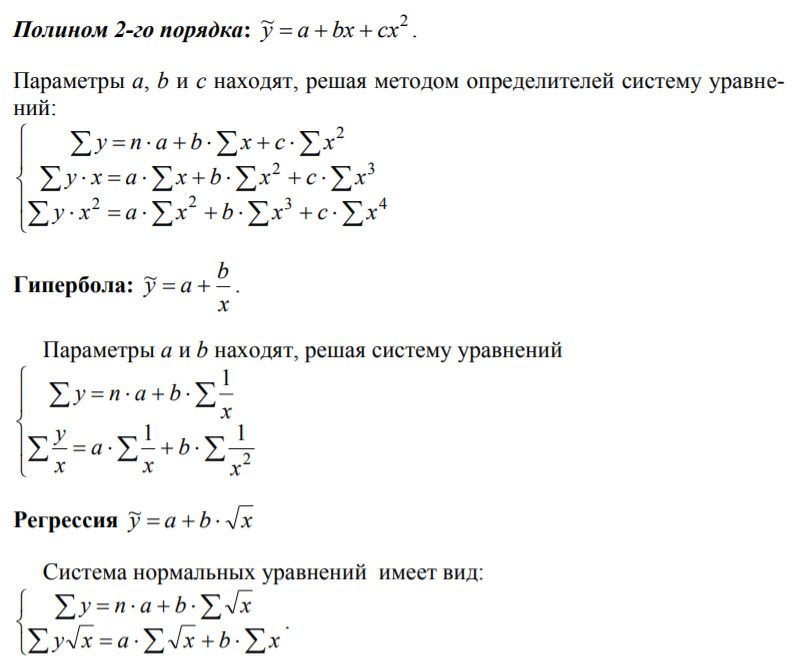
Проверил: Будникова И.К.

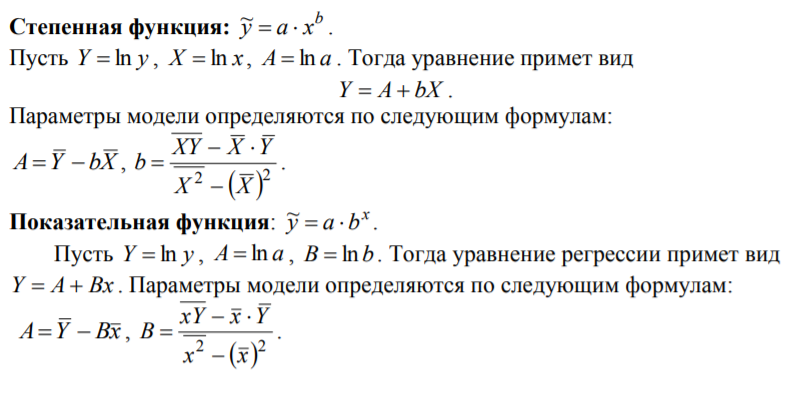
Казань 2024

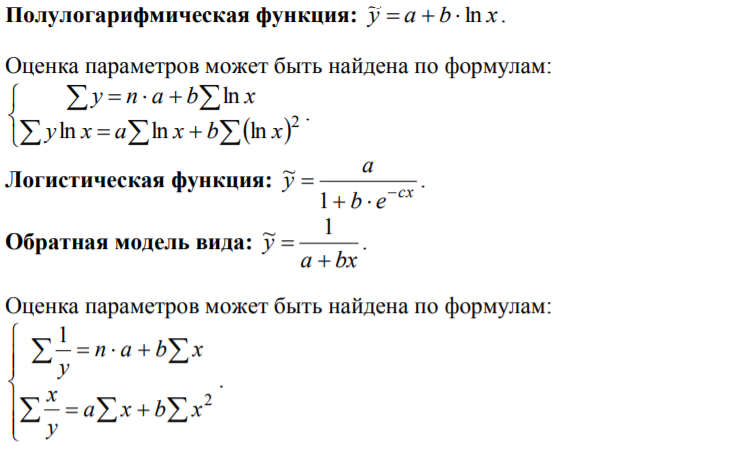
**Лабораторная работа № 2**

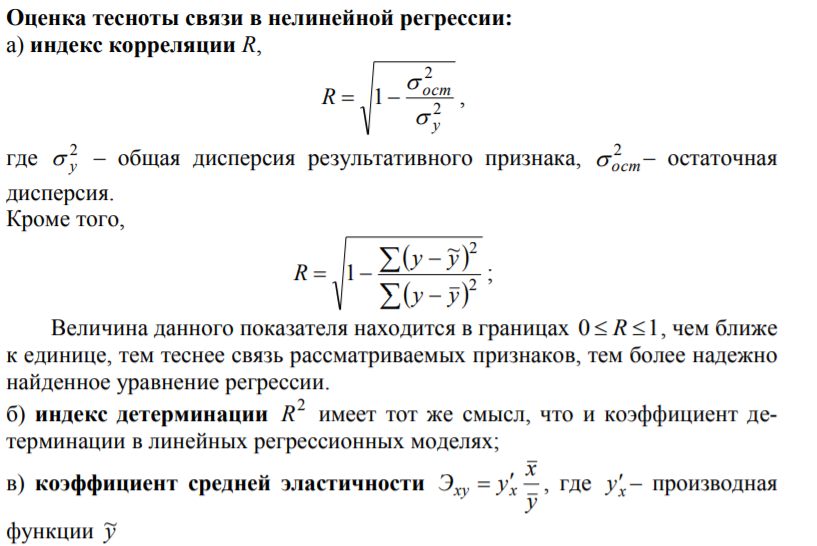
**Нелинейные модели парной регрессии**

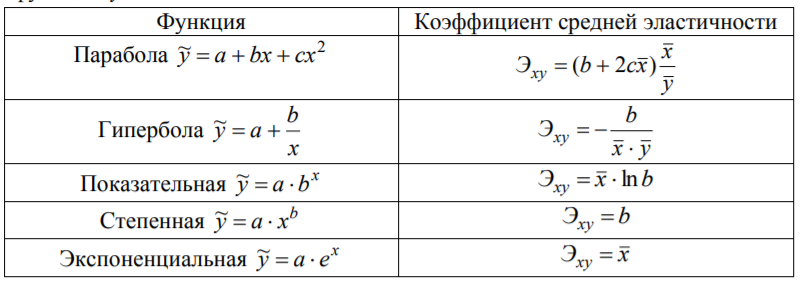
**Краткая теория**

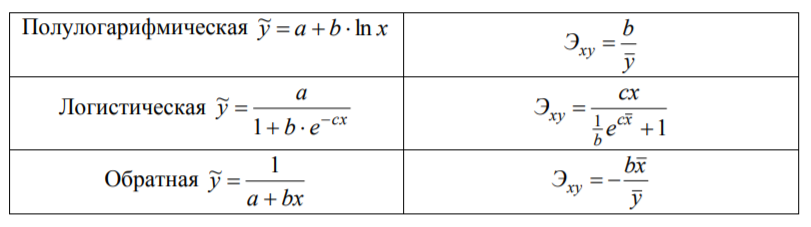
****

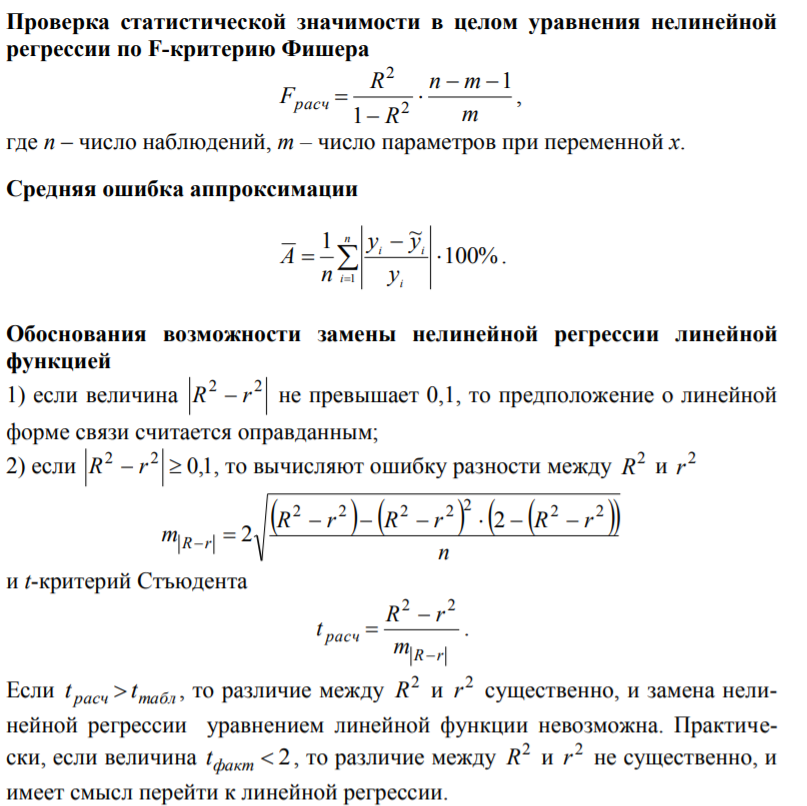
****

****

****

****

****

****

Задание на работу:

1. Построить поле корреляции и сформулировать гипотезу о форме связи.

2. Рассчитать параметры параболической, степенной, показательной, полу-

логарифмической, обратной и гиперболической регрессий.

3. Постройте на одной диаграмме с полем корреляции линию регрессии.

4. В каждом случае оцените тесноту связи с помощью показателей корреляции и детерминации.

5. Оценить с помощью средней ошибки аппроксимации качество модели.

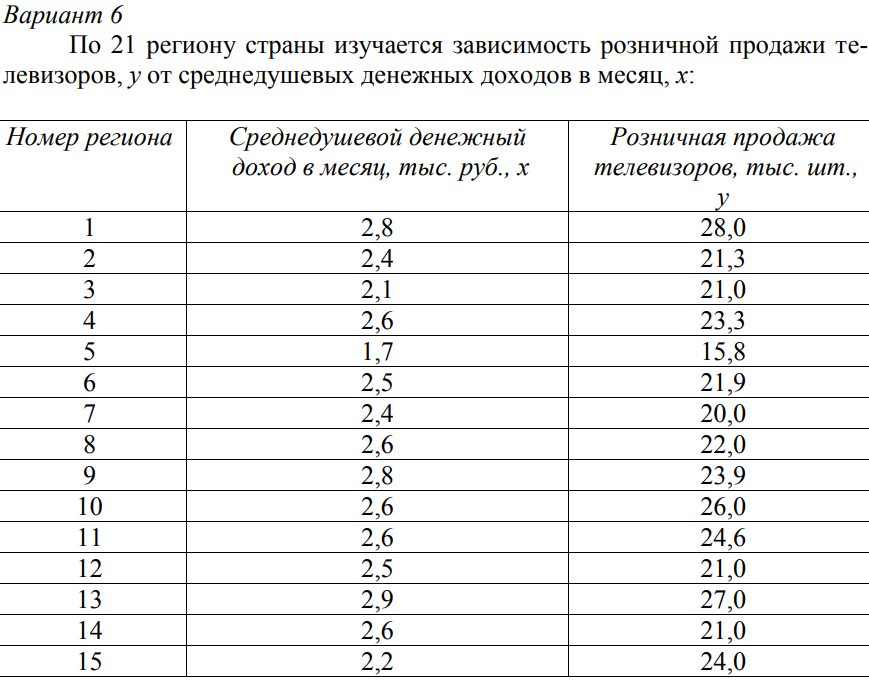
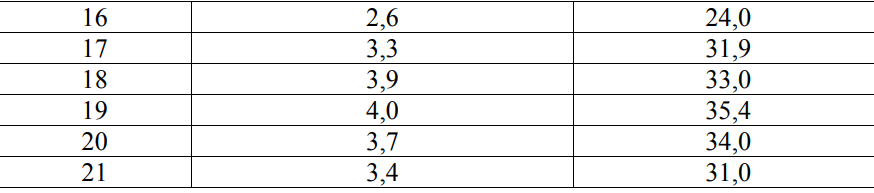
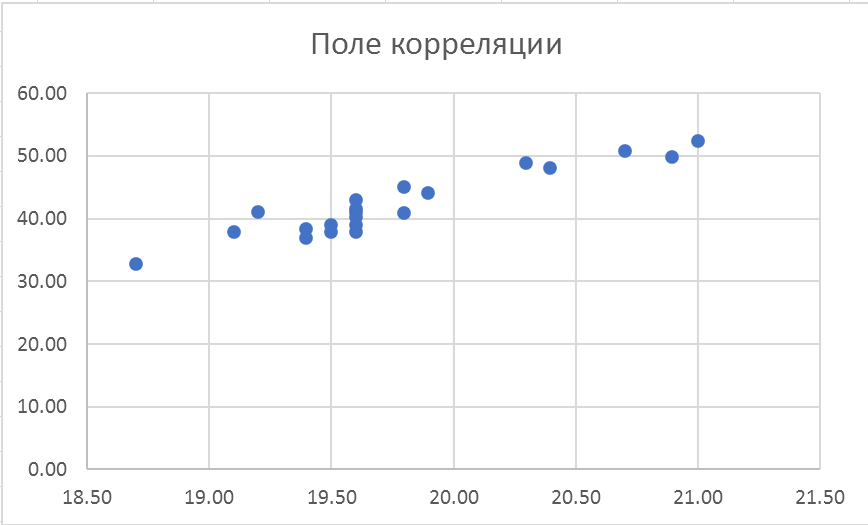
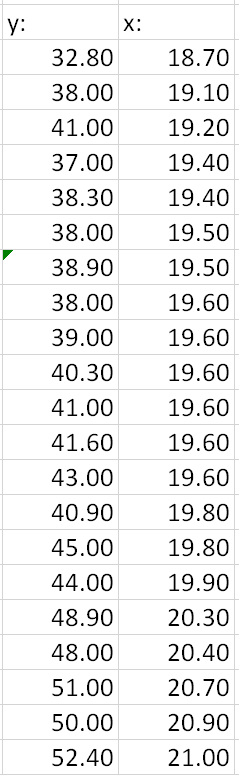
6. Оценить с помощью F-критерия Фишера статистическую надежность результатов регрессионного моделирования.

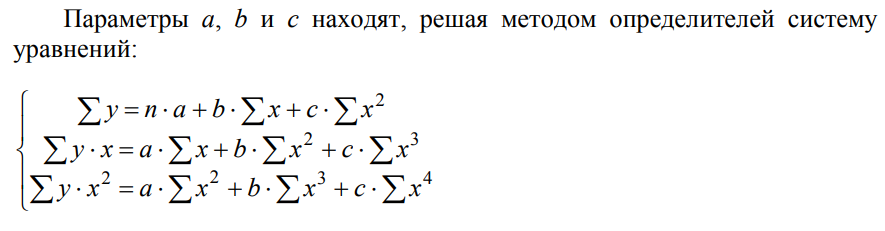
7. Выберите лучшее уравнение регрессии.

8. Дайте по выбранному уравнению оценку силы связи фактора с результатом с помощью среднего коэффициента эластичности.

9. Рассчитайте прогнозное значение результата, если прогнозное значение

фактора увеличится на 10% от его максимального в исходных данных значения. Определите доверительный интервал прогноза для уровня значимости альфа = 0,05.

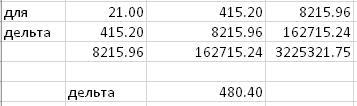
Исходные значения:  
  
  
  
**1. Построить поле корреляции и сформулировать гипотезу о форме связи.**

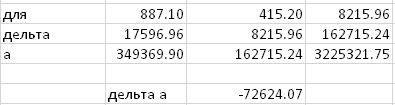
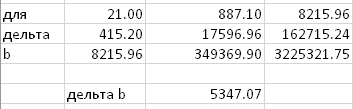
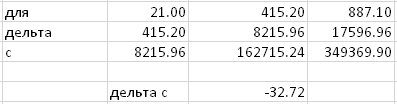
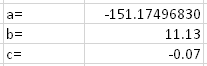
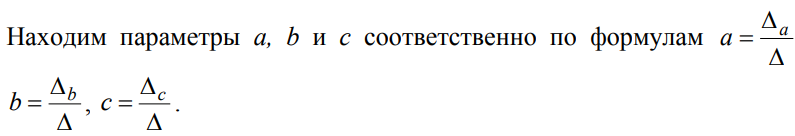
Форма связи близка к прямолинейной зависимости, т.к. точки расположены по прямой линии.  
  
  
  
**2. Рассчитать параметры параболической, степенной, полулогарифмической регрессий.  
Параболическая регрессия:**  


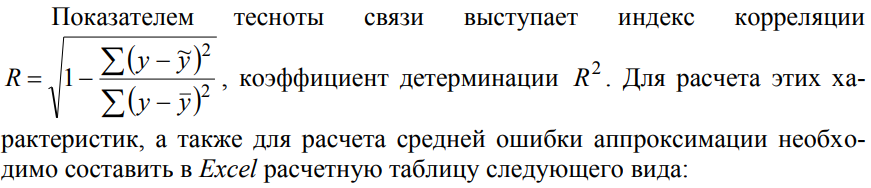
Получаем систему:  
21\*a + 415.2\*b + 8215.96\*c = 887.10

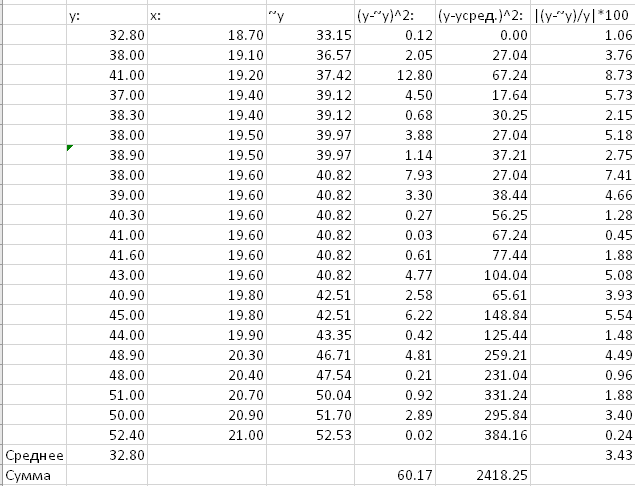
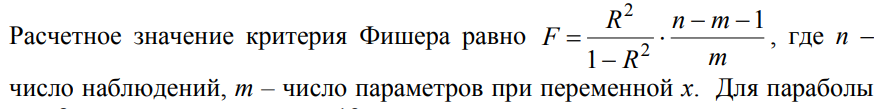
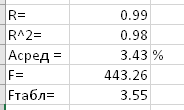
415.2\*a + 8215.96\*b + 162715.24\*c = 17596.96

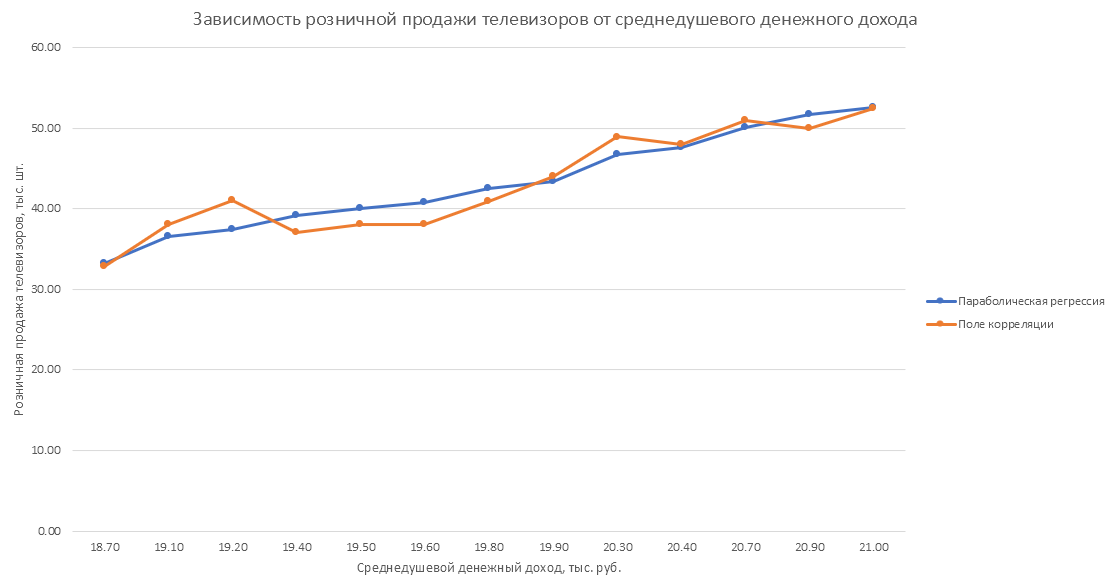
8215.96\*a + 162715.24\*b + 3225321.75\*c = 349369.9Составим и вычислим главный определитель системы, состоящий из коэффициентов при переменных a, b и c:

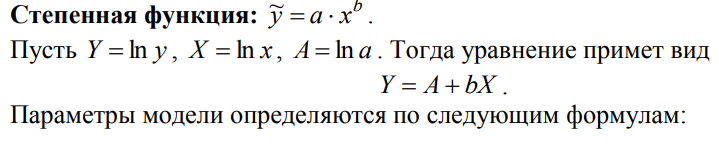
****Составляем и вычисляем три вспомогательных определителя системы:

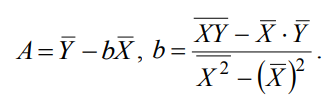
  
  
  
  
Таким образом уравнение регрессии принимает вид ~y = -0.07x2 + 11.13x – 151.17496830

****

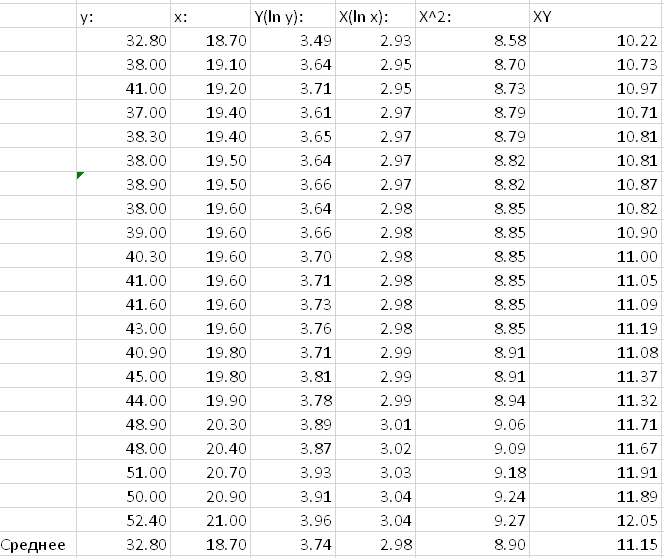
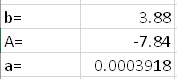
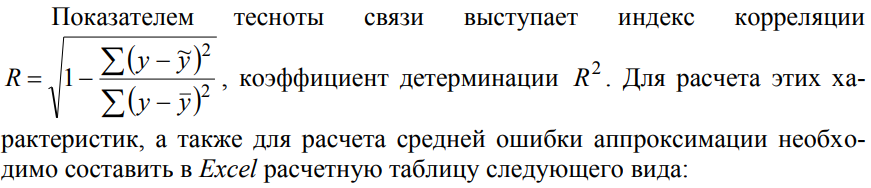
  
Тогда. R = sqrt(1 – (60.17 / 2418.25)) = 0.99, тогда R^2 = 0.98, Aсред. = 3.43%  
m = 2, в данном примере n = 21. F = 0.9801/(1 – 0.9801) \* (21 – 2 - 1) / 2 = 49.25 \* 9 = 443.26

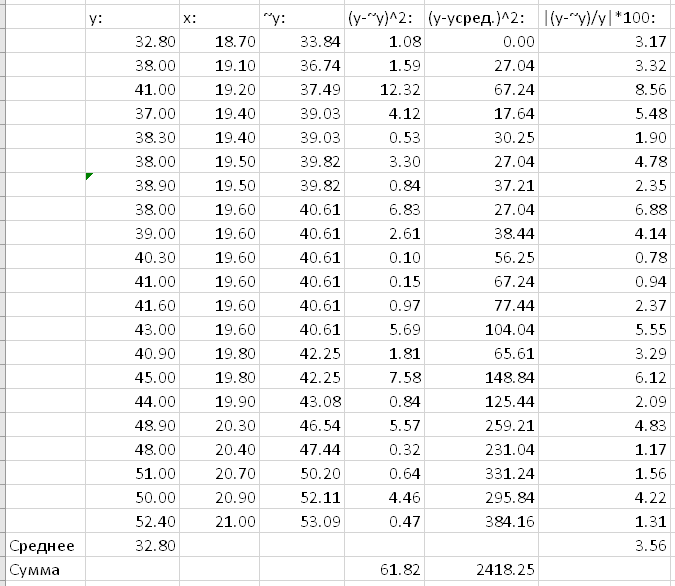
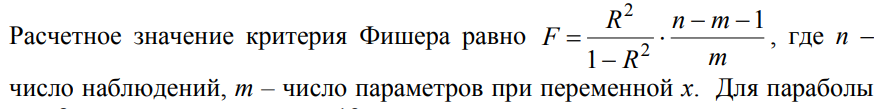
R = 0.99, что говорит о тесной прямой связи между признаками x и y.  
R2=0.98, т.е. 98% вариации признака y объясняется за счет признака x данным уравнением регрессии, что является весьма удовлетворительным  
Средняя ошибка аппроксимации не превышает установленного пределах в 15%, что свидетельствует о хорошем качестве модели.  
Расчётное значение критерия Фишера рано 443.26, оно превышает соответствующее табличное значение 3.55(=FРАСПОБР(0.05, 2**(2 фактора)**, 18**(21 – 2 - 1)**). Найденное уравнение параболической регрессии статистически надежно.  


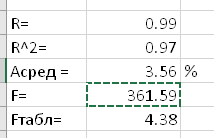
**Степенная регрессия:**



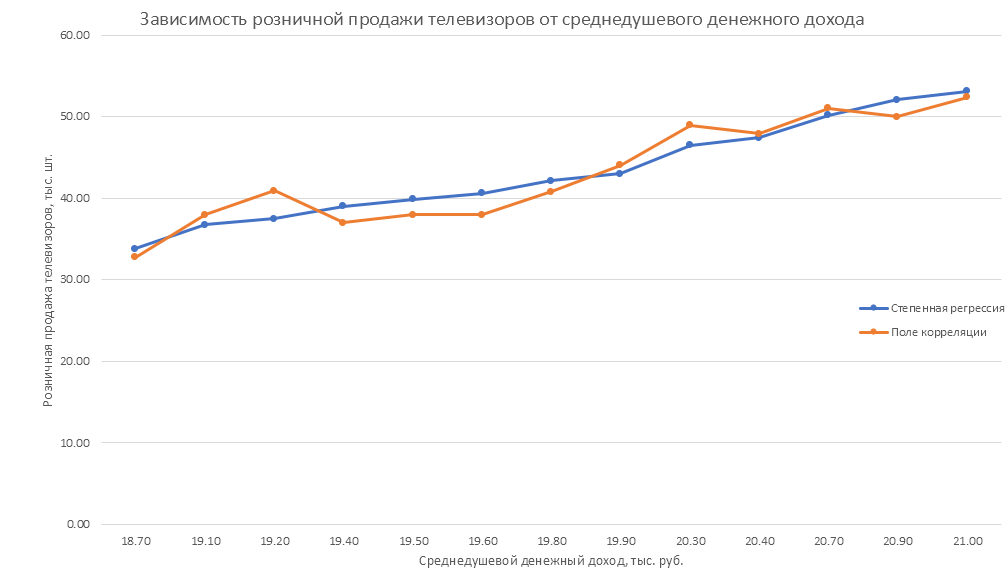
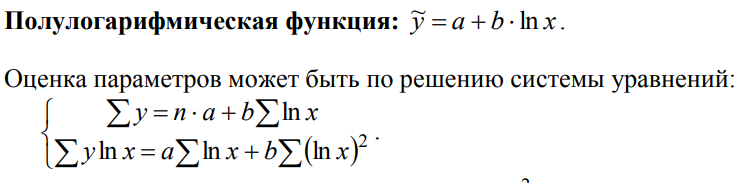
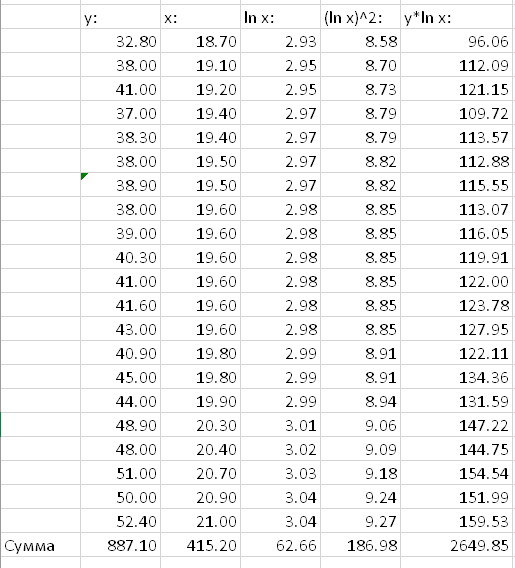
Составим вспомогательную таблицу:

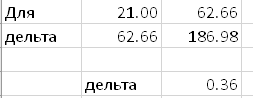
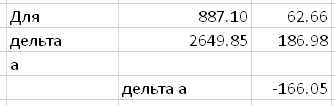
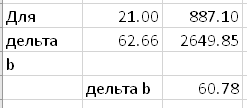
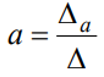
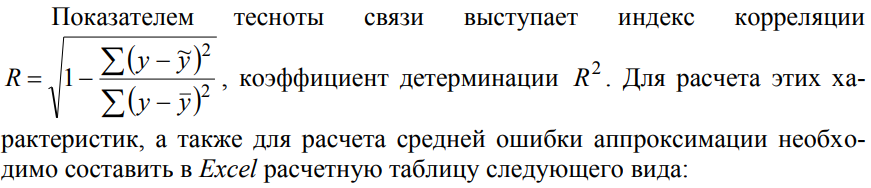
  
Тогда параметры будут равны:   
  
Уравнение регрессии примет вид: 0.00039\*x3.88****

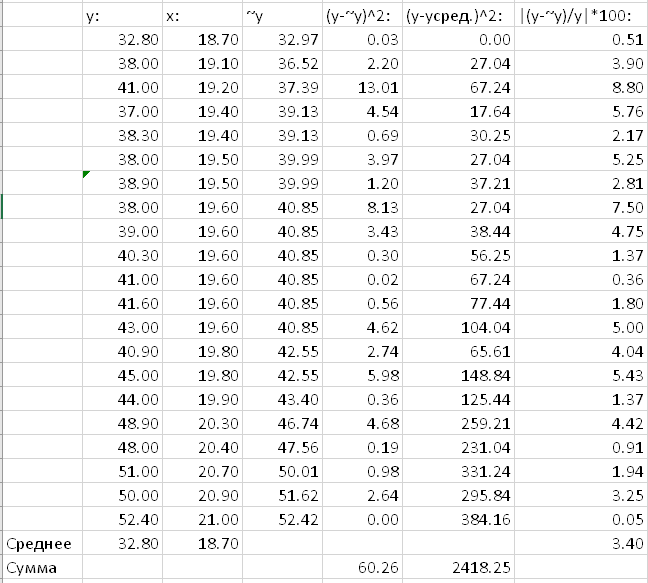
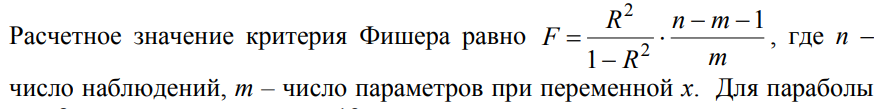
  
Тогда. R = sqrt(1 – (61.82 / 2418.25)) = 0.987, тогда R^2 = 0.9744, Aсред. = 3.56%  
m = 1, в данном примере n = 21. F = 0.9744/(1 – 0.9744) \* (21 – 1 - 1) / 2 = 38.06 \* 9.5 = 361.59

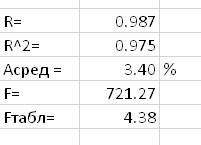


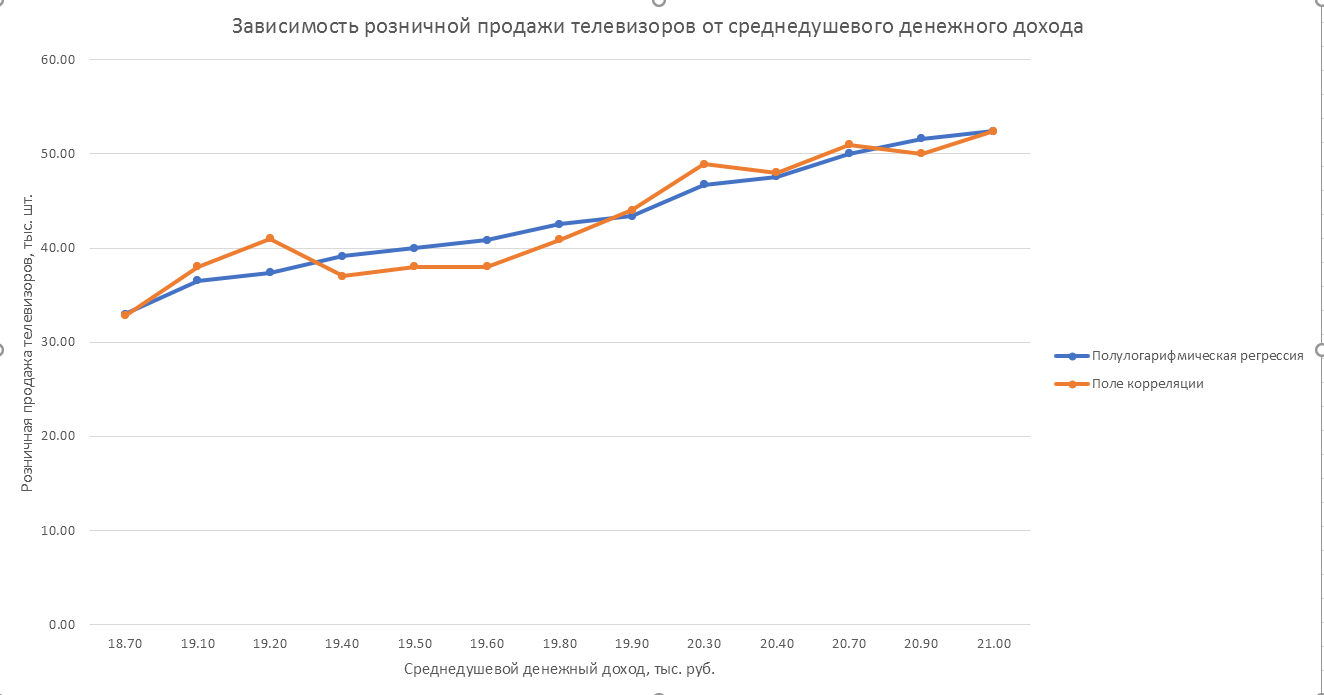
R = 0.987, что говорит о тесной прямой связи между признаками x и y.  
R2=0.97, т.е. 97% вариации признака y объясняется за счет признака x данным уравнением регрессии, что является весьма удовлетворительным  
Средняя ошибка аппроксимации не превышает установленного пределах в 15%, что свидетельствует о хорошем качестве модели.  
Расчётное значение критерия Фишера рано 361.59, оно превышает соответствующее табличное значение 4.38(=FРАСПОБР(0.05, 1**(1 фактор)**, 19**(21 – 1 - 1)**). Найденное уравнение параболической регрессии статистически надежно.

  
  
**Полулогарифмическая регрессия:**  
Составим вспомогательную таблицу:  
  
Получаем систему уравнений:  
21\*a + 62.66\*b = 887.1

62.66\*a + 186.98\*b = 2649.85  
Составим и вычислим основной определитель системы:  
  
Составим и вычислим вспомогательные определители системы:  
  
  
Находим параметры a и b соответственно по формулам ,  
  
Уравнение полулоарифмической регрессии примет вид:   
~y=167.74\*ln(x) – 458.25  
  
  
  
  
  
  
  
  
  
****

  
Тогда. R = sqrt(1 – (60.26 / 2418.25)) = 0.987, тогда R^2 = 0.975, Aсред. = 3.4%  
m = 1, в данном примере n = 21. F = 0.987/(1 – 0.987) \* (21 – 1 - 1) / 2 = 75.92 \* 9.5 = 721.26

Параметры полулогарифмической регрессии:  


R = 0.987, что говорит о тесной прямой связи между признаками x и y.  
R2=0.975, т.е. 97.5% вариации признака y объясняется за счет признака x данным уравнением регрессии, что является весьма удовлетворительным  
Средняя ошибка аппроксимации не превышает установленного пределах в 15%, что свидетельствует о хорошем качестве модели.  
Расчётное значение критерия Фишера рано 721.27, оно превышает соответствующее табличное значение 4.38(=FРАСПОБР(0.05, 1**(1 фактор)**, 19**(21 – 1 - 1)**). Найденное уравнение параболической регрессии статистически надежно.  


**7. Выберите лучшее уравнение регрессии.**  
Уравнение параболической регрессии: ~y = -0.07x2 + 11.13x – 151.17

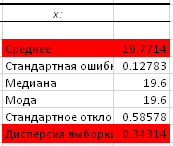
Уравнение степенной регрессии: ~y = 0.00039\*x3.88

Уравнение полулогарифмической регрессии: ~y=167.74\*ln(x) – 458.25

|  |  |  |
| --- | --- | --- |
| Регрессия | R2 | A cред. |
| Параболическая | 0.9751 | 3.43% |
| Степенная | 0.9744 | 3.56% |
| Полулогарифмической | 0.9751 | 3.40% |

Лучшим уравнением регрессии считается уравнение полулогаримической регрессии, т.к. оно обладает наименьшей средней ошибкой аппроксимации(3.4%)

**8. Дайте по выбранному уравнению оценку силы связи фактора с результатом с помощью среднего коэффициента эластичности.**  
Эxy = b/y = 167.74/42.24 = 3.97  
При увеличении среднедушевого денежного дохода на 1%, розничная продажа телевизоров увеличится ~3.97%  
  
**9. Рассчитайте прогнозное значение результата, если прогнозное значение фактора увеличится на 10% от его максимального в исходных данных значения. Определите доверительный интервал прогноза для уровня значимости α = 0,05.**Максимальное значение x в исходных данных = 21(тыс. руб.)  
  
Прогнозное значение х = 21 \* 1.1 = 23.1(тыс. руб)

Точечный прогноз признака y: 167.74\*ln(21) - 458.25 = 52.438(тыс. шт.)  
  
S2 = Dост. = 60.18175 / (21 - 2) = ~3.17  
  
S = 1.779736  


my = 1.78 \* sqrt(1 + 1/21 + (23.1 - 19.77)2/(21\*0.34)) = ~2.86  
  
Доверительный интервал прогноза:  
52.438 - 4.38 \* 2.86 <= yp <= 52.438 + 4.38 \* 2.86  
**39.90<= yp <= 64.97**

т.е. можно быть уверенным на 95%, что розничная продажа телевизоров при среднедушевом денежном доходе в 23.1 тыс. руб. будет варьироваться в указанных пределах(при точечном прогнозе выпуска готовой продукции в 52.438 тыс. шт.)