Name_____ Class_____

Regressions Packet

This packet is due IN CLASS on Monday April 8, 2013. If your class drops, the packet will be due on the following day in class.

Complete all of the questions in the packet and record your answers on this sheet. You must hand in the entire packet with this page attached on top.

Your packet will be graded and you will also take an in class quiz on this material.

Answers:

1. a) _____
b) _____
c) _____

2.a)_____
b) _____
c)_____

3. a) _____
b) _____
c) _____

4. a) _____
b) _____
c) _____

5. a) _____
b) _____
c) _____

6. a) _____
b) _____
c) _____

7. a) _____
b) _____
c) _____

8. a) _____
b) _____
c) _____

9. a) _____
b) _____
c) _____

10. a) _____
c) _____

11. a) _____
b) _____
c) _____

12. a) _____ r = _____
b) _____ r = _____
c) _____ r = _____
d) _____

13. a) _____ r = _____
b) _____ r = _____
c) _____ r = _____
d) _____

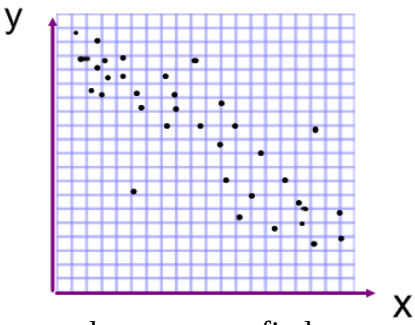
14. a) _____ r = _____
b) _____ r = _____
c) _____ r = _____
d) _____

15. a) _____ r = _____
b) _____ r = _____
c) _____ r = _____
d) _____

Regressions Packet This Packet is DUE April 16, 2012.
InClass. It will count as a quiz!

Scatterplots

A scatter plot graphically displays two related sets of data. Such a visual representation can indicate patterns, trends and relationships.

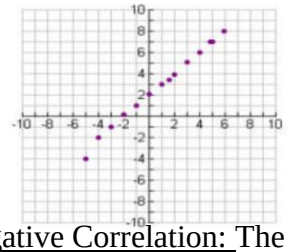


When we analyze the data from the scatterplot, we try to find an equation that can represent all of the data. Sometimes a line could be drawn to represent the points, other times a parabola could be drawn to represent the points, etc. We call the equations to the lines, parabolas, etc. **regression equations**. In this packet you will be calculating different regression equations. (Each problem will tell you which regression equation to calculate) Your calculator will find the equation for you.

After we find the equation to represent the points (data), we need to know how well the equation matches the data. For this we have a number called the **correlation coefficient**.

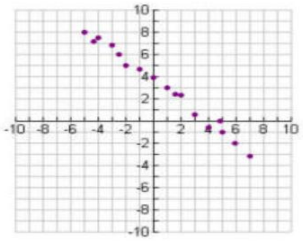
Correlations - The way the points are situated on the graph (represented by a number (r))

1. Positive Correlations: The points on the graph could represent a line with a positive slope



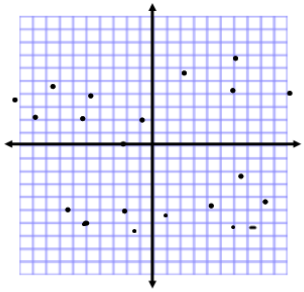
(r = 1)

2. Negative Correlation: The points on the graph could represent a line with a negative slope



(r = -1)

3. No Correlation:The points on the graph would not represent a line at all.



(r = 0)

Correlation Coefficient (r)

- Used to measure how well the data matches a line that would best fit the data.

-The correlation coefficient will be a number between -1 and 1.

**The closer the number is to 1 or -1, the better the data matches the line of best fit. The closer it is to zero, the worse the data matches the line of best fit.

Note: Negative does not mean it is a bad fit, it means the data slopes down.

- Your calculator will give you both r and the equation to match the data:

USING THE CALCULATOR FOR REGRESSIONS

1. **Clear the memory**

   press the [2nd] key then the [+] key then [7] (Reset)

   move the cursor to **ALL** and press [ENTER]

   press [2] : Reset

2. **Turn Diagnostic On**

   press the catalog button, [2nd] [0]

   scroll down to the D's and press [ENTER] next to DiagnosticOn

   the calculator will print DiagnosticOn on the screen, hit [ENTER] again and the calculator prints Done

3. **Enter the data**

   press the [STAT] key then [1] (Edit)

   Enter all x-values in L_1 (The first row of data)

   Enter all y-values in L_2 (The second row of data)

6. **Find the regression equation**

   press the [STAT] key

   move the cursor to **CALC**

   A) FOR LINEAR EQUATIONS: press [4] (LinReg(ax+b)) [ENTER]

   This will print on your screen:

   y = ax + b

   a: slope

   b: y-intercept

   r: correlation coefficient

   B) FOR QUADRATIC EQUATIONS: press 5 (QuadReg)

   C) FOR EXPOENTIAL EQUATIONS: press 0 (ExpReg)

7. **Write the regression equation**. (State as y = and plug all numbers into equation at top of screen)

8. **To find any other y-values**:

   substitute the x-values into the equation found in step 7

Regressions Packet

Linear Regressions

Model Problem:

A group of students set out to see if the hours of television they watched yesterday relates to their scores on today's test. The data is as follows:

Hours	0	0	0.5	1	1	1	1.5	2	2	2.5	2.5	3	3.5	4	4.5	5
Test Score	92	100	89	82	90	95	85	70	80	65	70	68	60	65	55	60

a) Find the **linear regression** equation that models this data (to the nearest hundredth)

y = -8.48x + 94.64

b) Find the correlation coefficient to 4 decimal places (rounded)

r = -.9293

c) Using the equation from part a, what would the grade be on the test if 6 hours of television was watched?

y = -8.48x +94.64 Ans: 44

y = -8.48(6) + 94.64

y = 43.76

y = 44

Complete the following problems:

1. The table below shows the number of fat grams and calories in some foods

Fat (grams)	8	6	4	19	14	1	10	1	9
Calories	150	80	90	245	270	70	185	45	115

a) Find the **linear regression** equation that models this data **(to the nearest tenth)**

_____

b) Find the correlation coefficient to 4 decimal places (rounded)

_____

c) Using the equation from part a, find the number of calories in a food that has 7 grams of fat. (nearest gram)

Ans: _____

2. Sarah is planning a party. Her favorite restaurant has given the cost for different amounts of guests.

# of guests	8	11	15	17	22	26	35	42	50
Cost	160	210	278	314	402	498	650	781	935

a) Find the **linear regression** equation that models this data **(to the nearest tenth)**

_____

b) Find the correlation coefficient to 4 decimal places (rounded)

_____

c) Using the equation from part a, find the cost of having 75 people at the party. (round to nearest person – for people, always round down)

Ans: _____

******Tables with years*******

Model Problem:

Greeting card Sales are recorded each year.

year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
sales(in billions)	2.05	2.3	2.45	2.7	3.2	3.45	3.65	3.75	3.9	4.2	4.6	5	5.35	5.6

****There is an extra row in the table to change the years to regular numbers*** This must be done first. Take the first year and make it 0. Then figure out the number for each year after that. Your table should now look like this:**

year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
sales(in billions)	2.05	2.3	2.45	2.7	3.2	3.45	3.65	3.75	3.9	4.2	4.6	5	5.35	5.6

****Now enter the data in the calculator, BUT use the new numbers for the years in List 1 and the sales in List 2. This is how <u>all tables that have years will be changed</u>.**

Now answer the following questions:

a) Find the **linear regression** equation that models this data (to the nearest hundredth)

y = .27x + 1.98

b) Find the correlation coefficient to 4 decimal places (rounded)

r = .9937

c) Find the amount of sales in 2000. (Remember you must change the year to a number before substituting)

year 2000 = year #20 y = .27x +1.98 Ans: 7.38 billion $

y = .27(20) + 1.98

y = 7.38

Complete the following problems:

3. The table below shows U.S. movie earnings

year	1990	1994	1998	2002	2004	2005	2006
Billions of $	38.1	49.7	61.4	78.6	85.1	89.8	93.5

*****Notice this table of years changes the years in different intervals throughout the table****

a) Find the **linear regression** equation that models this data **(to the nearest tenth)**

b) Find the correlation coefficient to 4 decimal places (rounded)

c) Using the equation found in part a), what might the movie earnings be in 2011? (remember change 2011 to a number first) Round to nearest tenth.

Ans: _____

Quadratic Regressions

Model Problem:

High School enrollment for East High was recorded for the years from 1950 to 1990 and is shown in the accompanying table. Years shown count from 1950, such that x = 0 corresponds to the year 1950 and x = 40 corresponds to 1990.

Year	0	5	10	15	20	25	30	35	40
Enrollment	460	365	330	318	302	314	362	403	485

Note: This table has already changed the years to numbers

a) Determine a **quadratic regression** equation to model this data (Round coefficients to nearest hundredth)

$y = .41x^2 - 15.56x + 448.85$

b) What is the correlation coefficient correct to 4 decimal places?

****Your calculator will only give your R^2, so you will have to take the square root****

$r = 0.9907$

c) Using the regression equation found in part b, estimate the enrollment of East High in 1973.

year 1973 = year # 23 $y = .41x^2 - 15.56x + 448.85$ Ans: 308 people

$y = .41(23)^2 - 15.56(23) + 448.85$

$y = 307.86$

Complete the following problems:

4. When debris is dropped from a cliff its height, h, measured in feet at various times, t, measured in seconds, is shown in the accompanying table.

Time	0	1	2	3	4	5
Height	875	843	807	741	662	450

a) Determine a **quadratic regression** equation to model this data **(Round coefficients to nearest tenth)**

_____

b) What is the correlation coefficient correct to 4 decimal places? (rounded)

_____

c) Using the regression equation, find the approximate height, to the nearest foot, of the debris at 7 seconds.

Ans: _____

5. The table below contains data modeled by a quadratic equation of the form $y = ax^2 + bx + c$.

x	1	5	8	11	13
y	-7	-19	14	83	149

Find a, b, and c.

a = _____

b = _____

c = _____

6. From 1995 to 1999, the yearly profits, P, of a local company are shown by year in the table below, where t= 5 represents 1995 and values for P are given in thousands of dollars.

t	5	6	7	8	9
P	750	860	980	1100	1280

a) Determine a **quadratic regression** equation to model this data

_____

b) What is the correlation coefficient correct to 4 decimal places? (rounded)

_____

c) Estimate the company's profits for 1993. (nearest dollar)

Ans:_____

7. The total sales, S, of TV antennas for various years from 1980 to 1995 are shown in the table below.

Years	1980	1985	1987	1993	1995
sales	76.3	82.2	84.6	80.9	77.3

a) Determine a **quadratic regression** equation to model this data **(Round coefficients to nearest tenth)**

_____

b) What is the correlation coefficient correct to 4 decimal places?

_____

c) Use the regression equation to predict total sales of antennas for 2008 (Round answer to the nearest tenth of a million)

Ans: _____

Exponential Regressions:

Model Problem:

Below is a table representing the growth of a town from 1997 to 2003.

Year		Population
1997	0	50,000
1998	1	54,000
1999	2	58,000
2000	3	62,986
2001	4	68,024
2002	5	73,466
2003	6	79,344

a. Find and write the model of an **exponential regression**. (tenth)
   $y = 49931.3(1.1)^x$
b. Find the correlation coefficient correct to 4 decimal places. (round to 4 places)
   $r = .9999$
c. Predict what the population will be in the year 2017.
   2017 = year # 20 $y = 49931.3(1.1)^x$ Ans: 335,913 people
   $y = 49931.3(1.1)^{20}$
   $y = 335,912.81$

Complete the following problems:

8) Estimates for world population vary, but the data in the accompanying table are reasonable estimates of the world population from 1800 to 2000.

Year	1800	1850	1900	1950	1970	1980	1990	2000
Population (millions)	980	1260	1650	2520	3700	4440	5270	6080

a. Generate a best fit **exponential function**. **(Round coefficients to nearest hundredth)**

_____

b. Find the correlation coefficient correct to 4 decimal places. (rounded)

_____

c. Using the function, estimate the world population in 2010 (in millions). (Nearest person – with people always round down)

Ans: _____

9) The accompanying table shows the annual mean personal income versus years of education

Education (years)	1	4	6	8	9	11	12	14	16	18	20
Income (thousands of $)	9	11	15	19	21	19	29	31	50	70	100

a. Find and write the model of an **exponential regression**. **(Round to nearest hundredth)**

_____

b. What is the correlation coefficient correct to 4 decimal places. (rounded)

_____

c. What would the income be (in thousands of dollars) for a person with 22 years of education? (round to the nearest dollar)

Ans: _____

10) The percentage of dentistry degrees awarded to women in the United State between 1970 and 1995 has gone up. The percent of male dentists from 1970 to 1995 is shown in the table below:

Year	1970	1975	1980	1985	1990	1995
% Men	36.4	30.9	20.7	13.3	3.1	0.9

a. Find an **exponential regression** model which best fits this data. **(Round to nearest hundredth)**

_____

c. What is the correlation coefficient correct to 4 decimal places? (rounded)

_____

11) The table below shows the number of chirps per minute that one cricket made as the temperature (in degrees Celsius) changed during a period of several hours.

Chirps per minute	105	105	125	125	130	149	153	152	164	171	175
Temperature	18	19	20	21	21	23	24	24	24	26	26

a) Write the **exponential regression** of best fit to the **nearest thousandth**.

_____

b) Find the correlation coefficient to four decimal places. (rounded)

_____

c) What would the temperature be if there were 87 chirps per minute? (nearest degree)

Ans: _____

BEST REGRESSION EQUATION

Directions: Find the best regression equation to fit the following tables of data.

To find the best regression equation:

1) Enter the data into the calculator
2) Find and state the regression equation and r for linreg, quadreg, expreg and lnreg
3) Your answer will be the equation with the r closest to 1 or -1. Write your final answer.

MODEL PROBLEM:

Eight students' SAT scores are recorded below:

Verbal	560	450	520	780	700	480	500	640
Math	680	520	560	750	680	540	550	700

ROUND ALL VALUES TO 3 DECIMAL PLACES

LinReg: Equation $y = .703x + 215.409$ r = .9229

QuadReg: Equation $y = -.002x^2 + 2.930x - 442.36$ r = .9463

ExpReg: Equation $y = 322.295(1.001)^x$ r = .9172

Final Answer: QuadReg: $y = -.002x^2 + 2.930x - 442.36$

Complete the following problems:

12. High School enrollment for West High was recorded for the years from 1950 to 1975 and is shown below:

Year	1950	1955	1960	1965	1970	1975
Enrollment	460	395	330	318	302	314

(Round to nearest hundredth)

a) LinReg: Equation _____ r = _____

b) QuadReg: Equation _____ r = _____

c) ExpReg: Equation _____ r = _____

d) Final Answer: _____

13. The win-loss record for the National League East teams shortly after the all-star break is shown in the accompanying table:

Team	W	L
Atlanta	58	35
Montreal	48	44
New York	46	46
Florida	45	47
Philadelphia	42	49

(Round to nearest hundredth)

a) LinReg: Equation _____ r = _____

b) QuadReg: Equation _____ r = _____

c) ExpReg: Equation _____ r = _____

d) Final Answer: _____

14. From 1990 to 1999, the yearly profits of a local company are shown in the table below:

Year	1990	1993	1997	1998	1999
Profit	750	860	980	1100	1280

(Round to nearest tenth)

a) LinReg: Equation _____ r = _____

b) QuadReg: Equation _____ r = _____

c) ExpReg: Equation _____ r = _____

d) Final Answer: _____

15. The blood pressure and weight of resting animals at the nearby animal Hospital were measured and recorded in the table below:

Weight	20	45	70	110	125	140
Blood Pressure	108	136	145	171	183	190

(Round to nearest thousandths)

a) LinReg: Equation _____ r = _____

b) QuadReg: Equation _____ r = _____

c) ExpReg: Equation _____ r = _____

d) Final Answer: _____