IEOR 165
Midterm 1 Sp12

1. **[5+5+5]** In a large study of the relationship between parental income and the IQs of their children, the following results were obtained:

   average income = $90,000, SD = $45,000
   average IQ = 100, SD = 15,
   r=0.5

   (a) For each income group ($0 − $9,999, $10,000 − $19,999, $20,000 − $29,999, *etc*), the average IQ of children with parental income in that group was calculated and then plotted above the midpoint of the group ($5,000, $15,000, $25,000, *etc*). It was found that the points on this graph followed a straight line very closely. What is the slope of the regression line?

   (b) What is an estimate of the income for a student with an IQ of 110?

   (c) Suppose we know that the parental income for a particular family is $120,000. What is an estimate of the child's IQ? Explain.

2. **[5+5+5+5+5+5+5+5]** Verbal and Math SAT scores are displayed in Figure 1. Scores are between 200 and 800 with no units. The summary statistics and regression output is given in the page following Figure 1 (labeled "SAT SCORES").

   (a) Describe the relationship in the scatter plot. Is there justification for a linear model? Explain.

   (b) Are there any unusual scores/outliers? Is there a justification, based on the information we have, for removing them? Comment.

   (c) What is the correlation coefficient? Interpret this statistic.

   (d) What is the $R^2$? Interpret this statistic.

   (e) What is the regression line? Interpret the slope of this line.

   (f) Predict the math score of a student with a verbal score of 650.

   (g) What is the standard deviation of the Math score, given that a student scores 650 for the verbal? Under the assumptions of the model, what will the standard deviation of the math score be if the verbal score was 700?

   (h) According to the model, what proportion of students scoring 650 in the verbal will have a math score above 700?

3. **[5+5]** A simple random sample of 1000 persons is taken to estimate the percentage of Democrats in a large population. It turns out that 543 of the people in the sample are Democrats.

   (a) Estimate the proportion of Democrats in the population and the standard error of the proportions estimate.

   (b) Calculate the 95% confidence interval for the proportion of Democrats in the population. Justify the assumptions that need to be satisfied for each step of this calculation.

4. **[15+5+5]** Figure 2 shows scatter plots and residual plots of 100 pairs of data. Specifically, the first pair shows $(x_i, y_i)$, the second $(x_i, log(y_i))$ and the third pair $(x_i, \sqrt{y_i})$. On the three pages following Figure 2, summary statistics for each of the variables and regression output is provided.

   (a) Using the provided output, build a model relating $y$ and $x$. Justify each of your steps.

   (b) Estimate $y$ given $x = 10$.

   (c) Given that $x = 10$, what is the probability that $y$ exceeds 15. Justify your approach.

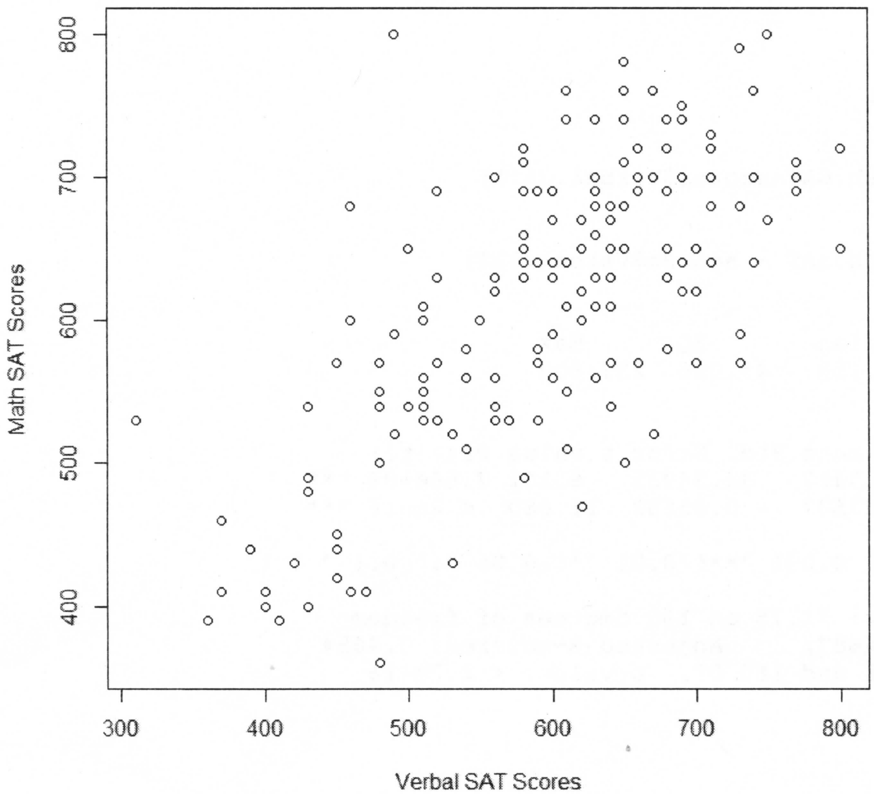


Figure 1: Scatter plot for math SAT scores against verbal SAT scores

```
> summary(scores)
   Verbal.SAT        Math.SAT      Sex
 Min.   :310.0   Min.   :360.0   F:82
 1st Qu.:522.5   1st Qu.:542.5   M:80
 Median :610.0   Median :630.0
 Mean   :596.3   Mean   :612.1
 3rd Qu.:660.0   3rd Qu.:690.0
 Max.   :800.0   Max.   :800.0
>
> sd(scores$Math.SAT)
[1] 98.13435
>
> sd(scores$Verbal.SAT)
[1] 99.5199
>
> summary(lm(scores$Math.SAT~scores$Verbal.SAT))

Call:
lm(formula = scores$Math.SAT ~ scores$Verbal.SAT)

Residuals:
     Min       1Q   Median       3Q      Max
-173.590  -47.596    1.158   45.086  259.659

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       209.55417   34.34935   6.101 7.66e-09 ***
scores$Verbal.SAT   0.67507    0.05682  11.880  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 71.75 on 160 degrees of freedom
Multiple R-squared: 0.4687,     Adjusted R-squared: 0.4654
F-statistic: 141.1 on 1 and 160 DF,  p-value: < 2.2e-16
```

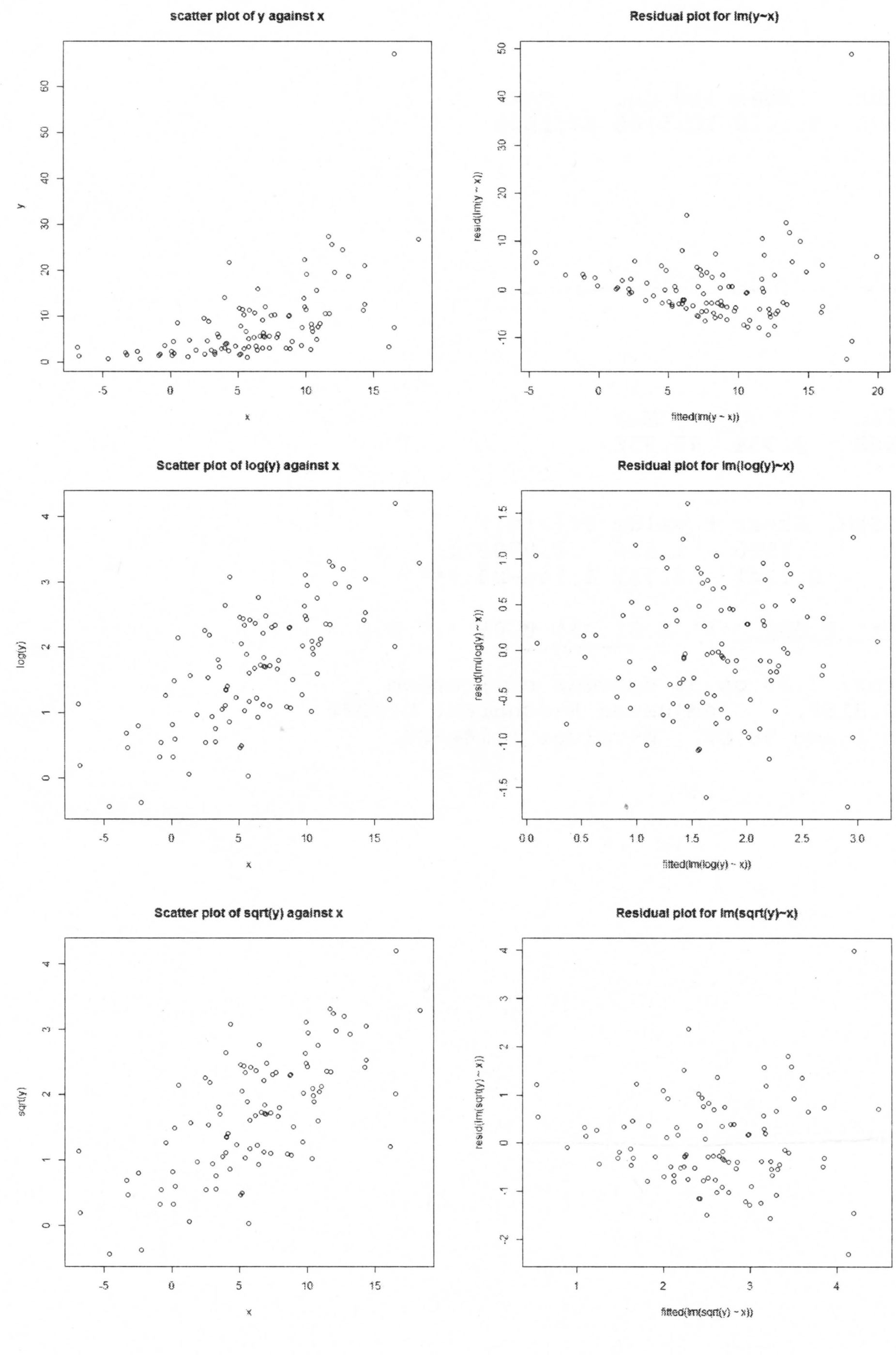


Figure 2: Scatter and residual plots

```
>
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 -6.900   3.406   6.284   6.224   9.901  18.320 
>
> sd(x)
[1] 5.042041
>
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 0.6422  2.8990  5.5170  8.1420 10.5100 67.1500 
>
> sd(y)
[1] 8.699141
>
> summary(lm(y~x))

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-14.431  -3.927  -0.868   2.984  48.958 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.1036     1.1520   1.826   0.0709 .  
x             0.9702     0.1441   6.732 1.14e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 7.23 on 98 degrees of freedom
Multiple R-squared: 0.3162,     Adjusted R-squared: 0.3092 
F-statistic: 45.32 on 1 and 98 DF,  p-value: 1.14e-09 
```

```
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -6.900   3.406   6.284   6.224   9.901  18.320
>
> sd(x)
[1] 5.042041
>
> summary(log(y))
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.4428  1.0640  1.7080  1.6960  2.3530  4.2070
>
> sd(log(y))
[1] 0.9083589
>
> summary(lm(log(y)~x))

Call:
lm(formula = log(y) ~ x)

Residuals:
     Min       1Q   Median       3Q      Max
-1.71081 -0.47647 -0.04866  0.46183  1.61413

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.93210    0.10647   8.754 6.08e-14 ***
x            0.12276    0.01332   9.217 6.06e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6682 on 98 degrees of freedom
Multiple R-squared: 0.4643,     Adjusted R-squared: 0.4589
F-statistic: 84.95 on 1 and 98 DF,  p-value: 6.056e-15
```

```
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -6.900   3.406   6.284   6.224   9.901  18.320
>
> sd(x)
[1] 5.042041
>
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.6422  2.8990  5.5170  8.1420 10.5100 67.1500
>
> sd(y)
[1] 8.699141
>
> summary(lm(y^0.5~x))

Call:
lm(formula = y^0.5 ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-2.3056 -0.5314 -0.2240  0.5460  3.9952

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.61440    0.14842  10.877  < 2e-16 ***
x            0.15589    0.01857   8.396  3.6e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9315 on 98 degrees of freedom
Multiple R-squared: 0.4184,     Adjusted R-squared: 0.4124
F-statistic: 70.49 on 1 and 98 DF,  p-value: 3.601e-13
```

# Cumulative Normal Probability Tables (Z-Values)

| Z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.50000 | 0.50399 | 0.50798 | 0.51197 | 0.51595 | 0.51994 | 0.52392 | 0.52790 | 0.53188 | 0.53586 |
| 0.1 | 0.53983 | 0.54380 | 0.54776 | 0.55172 | 0.55567 | 0.55962 | 0.56356 | 0.56749 | 0.57142 | 0.57535 |
| 0.2 | 0.57926 | 0.58317 | 0.58706 | 0.59095 | 0.59483 | 0.59871 | 0.60257 | 0.60642 | 0.61026 | 0.61409 |
| 0.3 | 0.61791 | 0.62172 | 0.62552 | 0.62930 | 0.63307 | 0.63683 | 0.64058 | 0.64431 | 0.64803 | 0.65173 |
| 0.4 | 0.65542 | 0.65910 | 0.66276 | 0.66640 | 0.67003 | 0.67364 | 0.67724 | 0.68082 | 0.68439 | 0.68793 |
| 0.5 | 0.69146 | 0.69497 | 0.69847 | 0.70194 | 0.70540 | 0.70884 | 0.71226 | 0.71566 | 0.71904 | 0.72240 |
| 0.6 | 0.72575 | 0.72907 | 0.73237 | 0.73565 | 0.73891 | 0.74215 | 0.74537 | 0.74857 | 0.75175 | 0.75490 |
| 0.7 | 0.75804 | 0.76115 | 0.76424 | 0.76730 | 0.77035 | 0.77337 | 0.77637 | 0.77935 | 0.78230 | 0.78524 |
| 0.8 | 0.78814 | 0.79103 | 0.79389 | 0.79673 | 0.79955 | 0.80234 | 0.80511 | 0.80785 | 0.81057 | 0.81327 |
| 0.9 | 0.81594 | 0.81859 | 0.82121 | 0.82381 | 0.82639 | 0.82894 | 0.83147 | 0.83398 | 0.83646 | 0.83891 |
| 1.0 | 0.84134 | 0.84375 | 0.84614 | 0.84849 | 0.85083 | 0.85314 | 0.85543 | 0.85769 | 0.85993 | 0.86214 |
| 1.1 | 0.86433 | 0.86650 | 0.86864 | 0.87076 | 0.87286 | 0.87493 | 0.87698 | 0.87900 | 0.88100 | 0.88298 |
| 1.2 | 0.88493 | 0.88686 | 0.88877 | 0.89065 | 0.89251 | 0.89435 | 0.89617 | 0.89796 | 0.89973 | 0.90147 |
| 1.3 | 0.90320 | 0.90490 | 0.90658 | 0.90824 | 0.90988 | 0.91149 | 0.91308 | 0.91466 | 0.91621 | 0.91774 |
| 1.4 | 0.91924 | 0.92073 | 0.92220 | 0.92364 | 0.92507 | 0.92647 | 0.92785 | 0.92922 | 0.93056 | 0.93189 |
| 1.5 | 0.93319 | 0.93448 | 0.93574 | 0.93699 | 0.93822 | 0.93943 | 0.94062 | 0.94179 | 0.94295 | 0.94408 |
| 1.6 | 0.94520 | 0.94630 | 0.94738 | 0.94845 | 0.94950 | 0.95053 | 0.95154 | 0.95254 | 0.95352 | 0.95449 |
| 1.7 | 0.95543 | 0.95637 | 0.95728 | 0.95818 | 0.95907 | 0.95994 | 0.96080 | 0.96164 | 0.96246 | 0.96327 |
| 1.8 | 0.96407 | 0.96485 | 0.96562 | 0.96638 | 0.96712 | 0.96784 | 0.96856 | 0.96926 | 0.96995 | 0.97062 |
| 1.9 | 0.97128 | 0.97193 | 0.97257 | 0.97320 | 0.97381 | 0.97441 | 0.97500 | 0.97558 | 0.97615 | 0.97670 |
| 2.0 | 0.97725 | 0.97778 | 0.97831 | 0.97882 | 0.97932 | 0.97982 | 0.98030 | 0.98077 | 0.98124 | 0.98169 |
| 2.1 | 0.98214 | 0.98257 | 0.98300 | 0.98341 | 0.98382 | 0.98422 | 0.98461 | 0.98500 | 0.98537 | 0.98574 |
| 2.2 | 0.98610 | 0.98645 | 0.98679 | 0.98713 | 0.98745 | 0.98778 | 0.98809 | 0.98840 | 0.98870 | 0.98899 |
| 2.3 | 0.98928 | 0.98956 | 0.98983 | 0.99010 | 0.99036 | 0.99061 | 0.99086 | 0.99111 | 0.99134 | 0.99158 |
| 2.4 | 0.99180 | 0.99202 | 0.99224 | 0.99245 | 0.99266 | 0.99286 | 0.99305 | 0.99324 | 0.99343 | 0.99361 |
| 2.5 | 0.99379 | 0.99396 | 0.99413 | 0.99430 | 0.99446 | 0.99461 | 0.99477 | 0.99492 | 0.99506 | 0.99520 |
| 2.6 | 0.99534 | 0.99547 | 0.99560 | 0.99573 | 0.99585 | 0.99598 | 0.99609 | 0.99621 | 0.99632 | 0.99643 |
| 2.7 | 0.99653 | 0.99664 | 0.99674 | 0.99683 | 0.99693 | 0.99702 | 0.99711 | 0.99720 | 0.99728 | 0.99736 |
| 2.8 | 0.99744 | 0.99752 | 0.99760 | 0.99767 | 0.99774 | 0.99781 | 0.99788 | 0.99795 | 0.99801 | 0.99807 |
| 2.9 | 0.99813 | 0.99819 | 0.99825 | 0.99831 | 0.99836 | 0.99841 | 0.99846 | 0.99851 | 0.99856 | 0.99861 |
| 3.0 | 0.99865 | 0.99869 | 0.99874 | 0.99878 | 0.99882 | 0.99886 | 0.99889 | 0.99893 | 0.99896 | 0.99900 |
| 3.1 | 0.99903 | 0.99906 | 0.99910 | 0.99913 | 0.99916 | 0.99918 | 0.99921 | 0.99924 | 0.99926 | 0.99929 |
| 3.2 | 0.99931 | 0.99934 | 0.99936 | 0.99938 | 0.99940 | 0.99942 | 0.99944 | 0.99946 | 0.99948 | 0.99950 |
| 3.3 | 0.99952 | 0.99953 | 0.99955 | 0.99957 | 0.99958 | 0.99960 | 0.99961 | 0.99962 | 0.99964 | 0.99965 |
| 3.4 | 0.99966 | 0.99968 | 0.99969 | 0.99970 | 0.99971 | 0.99972 | 0.99973 | 0.99974 | 0.99975 | 0.99976 |
| 3.5 | 0.99977 | 0.99978 | 0.99978 | 0.99979 | 0.99980 | 0.99981 | 0.99981 | 0.99982 | 0.99983 | 0.99983 |
| 3.6 | 0.99984 | 0.99985 | 0.99985 | 0.99986 | 0.99986 | 0.99987 | 0.99987 | 0.99988 | 0.99988 | 0.99989 |
| 3.7 | 0.99989 | 0.99990 | 0.99990 | 0.99990 | 0.99991 | 0.99991 | 0.99992 | 0.99992 | 0.99992 | 0.99992 |
| 3.8 | 0.99993 | 0.99993 | 0.99993 | 0.99994 | 0.99994 | 0.99994 | 0.99994 | 0.99995 | 0.99995 | 0.99995 |
| 3.9 | 0.99995 | 0.99995 | 0.99996 | 0.99996 | 0.99996 | 0.99996 | 0.99996 | 0.99996 | 0.99997 | 0.99997 |
| 4.0 | 0.99997 | 0.99997 | 0.99997 | 0.99997 | 0.99997 | 0.99997 | 0.99998 | 0.99998 | 0.99998 | 0.99998 |