Midterm Sample

STAT-UB 3 – Regression and Forecasting Models

Initials: ____________________

Please read this instruction carefully before the exam: The length of the exam is 60 min. The exam is closed book and notes, with the following exception: you are allowed to bring one letter-sized *one-sided* page of notes into the exam. You are also permitted use of a calculator (but smart phone is not allowed).

There are 8 multiple choices questions and 1 written questions. There are in total 80 points and 20 points are given for free (as long as you show up for the exam).

Please circle the choice which best answers each question *in the answer sheet* and and provide answers to written questions *in the answer sheet*. For both multiple choices questions and and written questions, only the answers on the answer sheet will be graded. *Answers not on the answer sheet will NOT be graded.* But you *have to turn in the whole booklet. The exam will be invalid if any page from this booklet is teared off.*

Additionally, it is your responsibility to make sure that the Teaching Fellow (TF) can clearly read your answer (if your circle for the multiple choice question is not clear enough or your handwriting is unreadable to TF, TF has the right to claim that this answer is wrong). Therefore, if you do need to modify your answer, make sure TF can easily recognize your final answer.

NYU Stern Honor Code:

I will not lie, cheat or steal to gain an academic advantage, or tolerate those who do.

Signature: ______________________________ Date: ____________________

Printed Name: ______________________________ Net ID: ____________________

Leave this page blank

Midterm Sample

STAT-UB 3 – Regression and Forecasting Models

Initials: ______________________

1)	(A)	(B)	(C)	(D)	(E)	**4)**	(A)	(B)	(C)	(D)	
2)	(A)	(B)	(C)	(D)	(E)	**5)**	(A)	(B)	(C)	(D)	(E)
3)	(A)	(B)	(C)	(D)	(E)	**6)**	(A)	(B)	(C)	(D)	(E)

Leave this page blank

Answer to Question 7.

Leave this page blank

Initials: ____________________

Answer to Question 8.

Leave this page blank

Midterm Sample
STAT-UB 3 – Regression and Forecasting Models
Initials: ____________________

Multiple Choice

1. (5 points) In simple linear regression if the correlation coefficient r is 0.7, then what proportion of variability in y is explained by x?
   A. 70%
   B. 84%
   C. 49%
   D. 100%
   E. None of the above

2. (5 points) In a regression data set, if r^2 is 0.75 and the total sum of squares is $SS_{yy} = 100$, then what is the value of the residual sum of squares, SSE?
   A. 25
   B. 75
   C. 100
   D. 175
   E. None of the above

3. (5 points) In a simple linear regression based on a very large sample (very large n), suppose that the t-statistic for the regression coefficients is 2.4. Then the two-tailed p-value corresponding to the given coefficient is (note that by the z-table, $P(0 < Z < 2.4) = 0.4918$ where $Z \sim \mathcal{N}(0, 1)$)
   - A. 0.9918
   - B. 0.0164
   - C. 0.9836
   - D. 0.0082
   - E. None of the above

4. (5 points) Consider two simple linear regression data sets, Set1 and Set2. If the sum of squared error SSE for Set2 is smaller than that for Set1, then
   - A. the r^2 for Set2 must be larger than the r^2 for Set1.
   - B. the r^2 for Set2 must be smaller than the r^2 for Set1.
   - C. the r^2 for Set2 must be the same as the r^2 for Set1.
   - D. None of the above.

5. (5 points) In simple linear regression, if the coefficient of determination if 0.5, then:
   - A. The estimated slope coefficient must be positive.
   - B. The estimated slope coefficient must be negative.
   - C. The data points must all lie on a straight line.
   - D. The sum of squares due to regression (SSR) must be less than the total sum of squares (SS_{yy}).
   - E. None of the above.

6. (5 points) In a simple linear regression with only one predictor variable x, if the sample size is 40, the total sum of squares SS_{yy} is 20, the regression sum of squares SSR is 5, what is the estimated standard error s:
   - A. 0.3947
   - B. 0.6283
   - C. 0.7500
   - D. 0.2500
   - E. None of the above.

7. (5 points) Retail price data on hard disk drives were recently reported in a computer magazine. The response variable y is the retail price (measured in a hundred dollars), the predictor x is microprocessor speed (measured in mega-hertz). The fitted simple linear regression model is $y = 3.73 + 1.05x$. Interpret the estimated slope for microprocessor speed (that is, $\hat{\beta}_1$)

   A. For every 1 mega-hertz increase in speed, the expected price increases 1.05 dollars.

   B. For every 1 dollar increase in price, the expected speed increases 1.05 mega-hertz.

   C. For every 1 mega-hertz increase in speed, the expected price increases 105 dollars.

   D. For every 1 dollar increase in price, the expected speed increases 105 mega-hertz.

   E. None of the above.

8. (5 points) Consider the response variable y and two possible predictors of y – variables p and q. Plot A shows the regression of y on p and plot B shows the regression of y on q. Let $\hat{\epsilon}_p$ denote the residuals from the regression of y on p and let $\hat{\epsilon}_q$ denote the residuals from the regression of y on q. Which of the following conclusions is true?

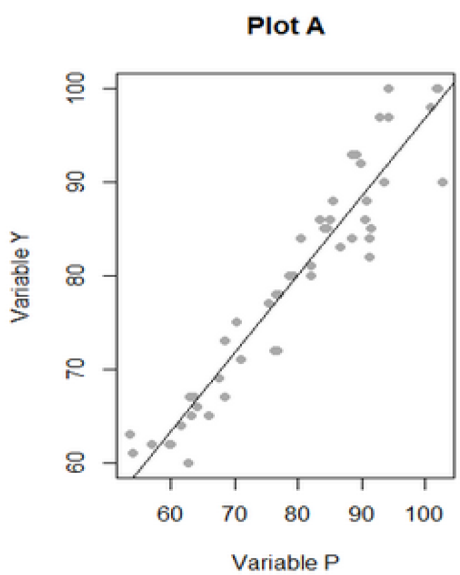
   A. p is the better predictor of y since $Var(\hat{\epsilon}_p)$ is greater than $Var(\hat{\epsilon}_q)$.

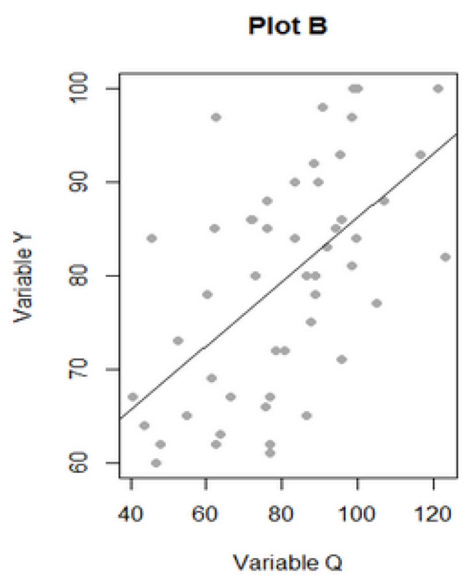
   B. p is the better predictor of y since $Var(\hat{\epsilon}_p)$ is less than $Var(\hat{\epsilon}_q)$.

   C. q is the better predictor of y since the regression of y on q has the greater r^2.

   D. q is the better predictor of y since the regression of y on q has the smaller r^2.

   E. None of the above.





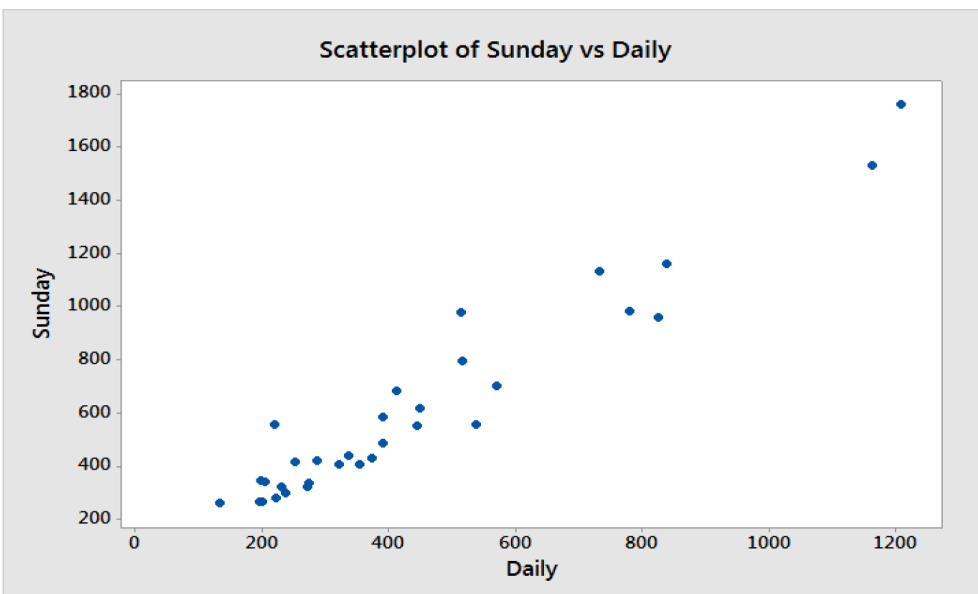
Written Problem

Important note: please write down your answers in the answer sheet (not here!). In addition to the final numerical solution in the answer sheet, please provide all the detailed intermediate steps and formulas in the answer sheet.

9. (40 points) In order to investigate the feasibility of starting a Sunday edition for a large metropolitan newspaper, information was obtained from a sample of newspapers concerning their Daily and Sunday circulations (both in thousands). In the Minitab output below Daily denotes daily circulation and Sunday, a dependent variable, denotes Sunday circulation. The basic descriptive statistics, scatter plot and the regression analysis output from the Minitab are shown below (some numbers are concealed and denoted by (1), (2), . . . ,etc).

```
Descriptive Statistics: Sunday, Daily

Variable   N   Mean  StDev  Minimum  Maximum
Sunday    32  614.4  375.9    262.0   1762.0
Daily     32  441.0  273.7    133.2   1209.2
```



```
Regression Analysis: Sunday versus Daily

The regression equation is
Sunday = 30.67 + 1.324 Daily


S = (2)   R-Sq =(3)


Analysis of Variance

Source       DF        SS
Regression    1   4070375
Error        30    310843
Total        31       (4)

Coefficients

Term        Coef  SE Coef  T-Value  P-Value
Constant   30.67     34.5     0.89    0.381
Daily      1.324   0.0668      (5)      (6)
```

(a) (10 points) Compute the t-value in (5) (round to the 2nd decimal place) using the MINITAB output. Is the regression of Sunday circulation on daily circulation *statistically significant* at $\alpha = 0.05$? Note you need to state (1) the hypothesis test, (2) rejection region, and (3) your conclusion.

(b) (5 points) Compute the total sum of squares $SS_{yy} = \sum_{i=1}^{n}(y_i - \overline{y})^2$ in (4) using the MINITAB output.

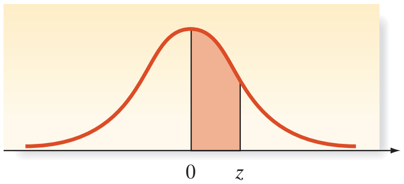
(c) (5 points) Compute the estimated standard error of the regression model s in (2) using the MINITAB output (round to the 3rd decimal place).

(d) (5 points) Compute the coefficient of determination (r^2) in (1) using the result from the precious question (round to the 3rd decimal place). Interpret the value of r^2 you obtained.

(e) (10 points) One is interested in the expected Sunday circulation when daily circulation is 400 (thousand). Calculate the fitted value and construct 95% confidence interval (round to the 1st decimal place).

(f) (5 points) Construct 95% prediction interval for sales when daily circulation is 400 (thousand) (round to the 1st decimal place).

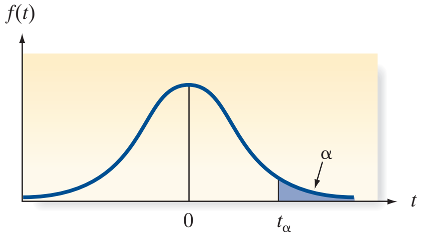
Table II Normal Curve Areas



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.49903	.49906	.49910	.49913	.49916	.49918	.49921	.49924	.49926	.48829
3.2	.49931	.49934	.49936	.49938	.49940	.49942	.49944	.49946	.49948	.49950
3.3	.49952	.49953	.49955	.49957	.49958	.49960	.49961	.49962	.49964	.49965
3.4	.49966	.49968	.49969	.49970	.49971	.49972	.49973	.49974	.49975	.49976
3.5	.49977	.49978	.49978	.49979	.49980	.49981	.49981	.49982	.49983	.49983
3.6	.49984	.49985	.49985	.49986	.49986	.49987	.49987	.49988	.49988	.49989
3.7	.49989	.49990	.49990	.49990	.49991	.49991	.49992	.49992	.49992	.49992
3.8	.49993	.49993	.49993	.49994	.49994	.49994	.49994	.49995	.49995	.49995
3.9	.49995	.49995	.49996	.49996	.49996	.49996	.49996	.49996	.49997	.49997

Leave this page blank

Table III **Critical Values of *t***



Degrees of Freedom	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	1.291	1.662	1.987	2.369	2.632	3.183	3.402
100	1.290	1.660	1.984	2.364	2.629	3.174	3.390
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
150	1.287	1.655	1.976	2.351	2.609	3.145	3.357
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Leave this page blank

Use this page as scratch paper only. Results on this page will not be counted for any credit.

Use this page as scratch paper only. Results on this page will not be counted for any credit.