

Statistics 2000, Section 001, Midterm 1 (200 Points)

Friday, September 25, 2009

Your Name: _____

Question 1: z-Scores and Normal Distributions (50 Points)

There are two major tests for readiness for college, the ACT and the SAT. ACT scores are reported on a scale from 1 to 36. The distribution of ACT scores for more than 1 million students in a recent high school graduating class was roughly Normal with mean $\mu = 20.8$ and standard deviation $\sigma = 4.8$. SAT scores are reported on a scale from 400 to 1600. The distribution of SAT scores for 1.4 million students in the same graduating class was roughly Normal with mean $\mu = 1026$ and standard deviation $\sigma = 209$.

Show your work!

1. (10 Points) Compare a SAT with an ACT score: Wendy scores 1350 on the SAT. Jeremy scores 25 on the ACT. Assuming that both tests measure the same thing, who has the higher score — Wendy or Jeremy? Report the z-scores for both students.

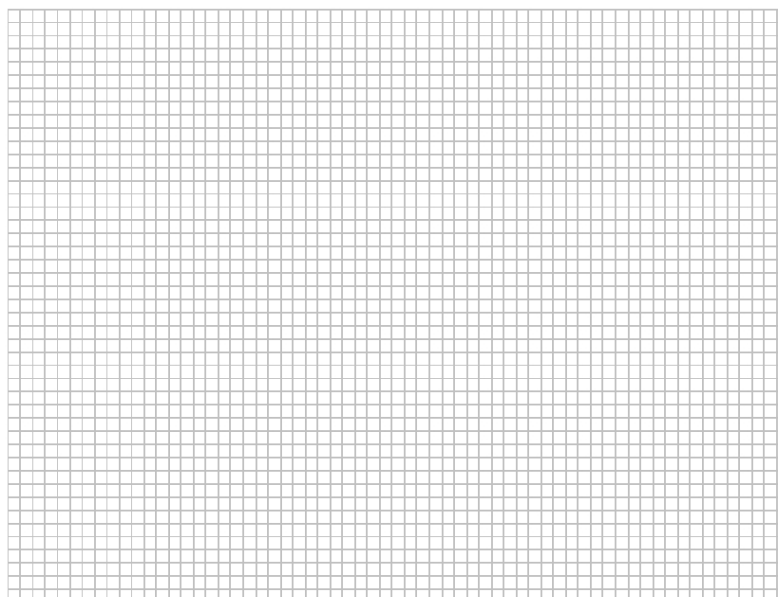
2. (10 Points) Find the ACT equivalent: Rob scores 1420 on the SAT. Assuming that both tests measure the same thing, what score on the ACT is equivalent to Rob's SAT score?

Question 2: Histograms (40 Points)

Fish in the Bering Sea. “Recruitment”, the addition of new members to a fish population, is an important measure of the health of ocean ecosystems. Here are data on the recruitment of rock sole in the Bering Sea between 1973 and 2000:

Year	Recruitment (millions)	Year	Recruitment (millions)
1973	173	1987	4700
1974	234	1988	1702
1975	616	1989	1119
1976	344	1990	2407
1977	515	1991	1049
1978	576	1992	505
1979	727	1993	998
1980	1411	1994	505
1981	1431	1995	304
1982	1250	1996	425
1983	2246	1997	214
1984	1793	1998	385
1985	1793	1999	445
1986	2809	2000	676

1. **(30 Points)** Draw a histogram to display the distribution of rock sole recruitment. You should work with 5 classes. Use the graph paper provided below. Make sure to label the axes.



2. **(10 Points)** Describe the pattern you see in your histogram and mention any striking deviations that you see.

Question 3: CrunchIt Output (50 Points)

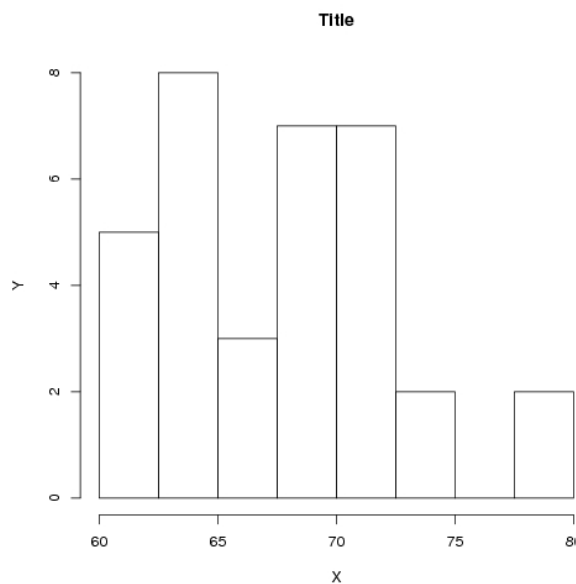
Shown below is output from CrunchIt based on Height and Weight of our Stat 2000 student data. Only 34 complete records have been considered.

Columns -- Selected Field: WEIGHT

Statistic	Result
n	34.0000
mean	159.8824
variance	1225.6627
std.dev	35.0098
std.err	6.0041
median	162.5000
range	180.0000
min	105.0000
max	285.0000
q1	130.0000
q3	180.0000
mode	no unique mode

Columns -- Selected Field: HEIGHT

Statistic	Result
n	34.0000
mean	67.9853
variance	21.8861
std.dev	4.6783
std.err	0.8023
median	68.0000
range	19.0000
min	61.0000
max	80.0000
q1	64.0000
q3	71.6250
mode	no unique mode



Linear Regression -- Select Fields: WEIGHT, HEIGHT

Estimates	Estimate	t value	Pr(> t)	Std. Error	
WEIGHT	0.0898	5.1302	0.0000	0.0175	
(Intercept)	53.6328	18.7397	8.1611e-19	2.8620	
ANOVA	Df	Sum Sq	Mean Sq	F value	Pr(>F)
WEIGHT	1.0000	325.9466	325.9466	26.3194	0.0000
Residuals	32.0000	396.2961	12.3843	NaN	NaN
R.sq	0.4513				
adjusted R.sq	0.4342				
F value	26.3194				

Correlation -- Select Fields: WEIGHT, HEIGHT

	WEIGHT	HEIGHT
WEIGHT	1.0000	0.6718
HEIGHT	0.6718	1.0000

Answer the questions on the next page based on the output provided above.

1. **(10 Points)** List the values for the five number summary for Height and clearly name them (e.g., if variance is one of these numbers than indicate “variance: ...”).

1.)

2.)

3.)

4.)

5.)

2. **(10 Points)** Read off the values (as stated in the CrunchIt output) that allow to predict Height (response) from Weight (explanatory) and combine them in the regression equation:

$$\text{predicted Height} = \text{-----} + \text{-----} * \text{Weight}$$

3. **(15 Points)** Manually calculate the values that allow to predict Weight (response) from Height (explanatory).

Calculation:

Then combine them in the regression equation:

$$\text{predicted Weight} = \text{-----} + \text{-----} * \text{Height}$$

4. **(5 Points)** Predict the weight for someone who is 72 inches tall. Indicate which equation you use (either from 2. or 3.) and write down your calculation and the final result.

5. **(10 Points)** One histogram is shown. Unfortunately, it is not labeled. Based on all other information, does this histogram relate to (i) Weight, (ii) Height, or (iii) none of these two variables? Circle your answer and explain.

Question 4: Multiple Choice Questions (60 Points)

Circle your answer. There is only one correct answer for each question. Each correct answer is worth 4 points.

1. Data on the mileage of 20 randomly selected cars are listed below. The values are ordered for convenience.

12 13 15 16 16 17 18 18 19 19
20 20 22 23 24 26 26 27 27 29

What is the median mileage for these 20 cars?

- (a) 17.5 miles per gallon
 - (b) 19 miles per gallon
 - (c) 19.5 miles per gallon
 - (d) 20 miles per gallon
2. A biology class has 201 students. The five number summary for the midterm exam was

23, 51, 62, 78, 92

The student with the 92 found a grading error on her exam and her correct grade was 95. There were no other grading errors. After correcting this student's paper, the five number summary for the midterm exam will be

- (a) 23, 51, 61.8, 78, 95.
 - (b) 23, 51, 62, 81, 95.
 - (c) 23, 51, 62, 78, 92.
 - (d) 23, 51, 62, 78, 95.
3. When ordering vinyl replacement windows, the following variables are specified for each window. Which of these variables is quantitative?
- (a) Window style — double hung, casement, or awning.
 - (b) Area of the window opening in square inches.
 - (c) Window style — single pane or double pane.
 - (d) Window manufacturer.

4. At a large department store, the amount a shopper spent and the shopper's gender (male or female) were recorded. To determine if gender is useful in explaining the amount of money a shopper spends at the store we could
 - (a) make side-by-side boxplots of the distribution of the amount spent by males and the distribution of the amount spent by females.
 - (b) compute the correlation between the amount spent and gender.
 - (c) compute the least-squares regression line of amount spent on gender.
 - (d) create a scatterplot of amount spent (response variable) and gender (explanatory variable).

5. Determine which of the following statements regarding the correlation coefficient is true. There is just one true statement.
 - (a) The correlation coefficient equals the proportion of times that two variables lie on a straight line.
 - (b) The correlation coefficient will be $+1.0$ if all the data points lie on a perfectly horizontal straight line.
 - (c) The correlation coefficient measures the strength of any relationship that may be present between two variables.
 - (d) The correlation coefficient is a unitless number and must always lie between -1.0 and $+1.0$, inclusive.

6. A company has 20 female workers whose average salary is \$43,000 and 30 male workers whose average salary is \$47,000. What is true about the average salary of all 50 workers?
 - (a) It must be \$45,000.
 - (b) It must be \$45,400.
 - (c) It must be \$47,000 because we have more male workers than female workers.
 - (d) It could be any number between \$43,000 and \$47,000.

7. A sample was taken of the salaries of 20 employees from a large company. The following are the salaries (in thousands of dollars) for this year (the data are ordered).

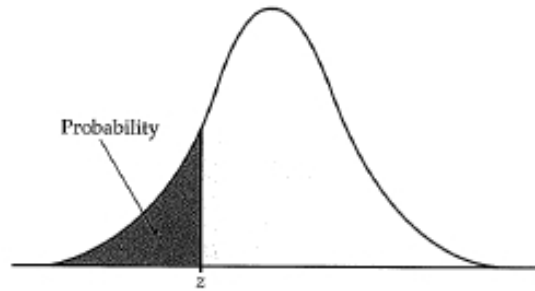
28 31 34 35 37 41 42 42 42 47
 49 51 52 52 60 61 67 72 75 77

Suppose each employee in the company receives a \$3,000 raise for next year (each employee's salary is increased by \$3,000). The interquartile range of the salaries will

- (a) be unchanged.
- (b) increase by \$3,000.
- (c) be multiplied by 3000.
- (d) become 50 (i.e., \$50,000).

8. A student group investigates the relationship between IQ and GPA, measured on a 12-point scale. They find the equation to be $GPA = -6 + 0.15 \cdot IQ$. Along comes Marilyn Vos Savant with an IQ of 200. What does this regression say her GPA should be?
- (a) 12
 - (b) 24
 - (c) 36
 - (d) Using the line is meaningless for her.
9. Researchers are conducting a state-wide survey for the U.S. Postal Service. The survey records many different variables of interest. Which of the following variables is categorical? There is just one.
- (a) County of residence.
 - (b) Number of people, both adults and children, living in the household.
 - (c) Total household income, before taxes, in 2003.
 - (d) Age of respondent.
10. In a statistics course, a linear regression equation was computed to predict the final exam score from the score on the midterm exam. The equation of the least-squares regression line was $y = 10 + 0.9 \cdot x$, where y represents the final exam score and x is the midterm exam score. Suppose Joe scores a 90 on the midterm exam. What would be the predicted value of his score on the final exam?
- (a) 81
 - (b) 89
 - (c) 91
 - (d) Cannot be determined from the information given. We also need to know the correlation.
11. All but one of the following statements contain a blunder. Which one does not contain a blunder?
- (a) "There is a correlation of $r = 0.54$ between the position a football player plays and his or her weight."
 - (b) "The correlation between amount of fertilizer and yield of tomatoes was found to be $r = 0.33$."
 - (c) "The correlation between the gas mileage of a car and its weight is $r = 0.71$ gallon-pounds."
 - (d) "The correlation between amount of fertilizer and yield of tomatoes was found to be $r = 2.00$."

Table entry for z is the area under the standard normal curve to the left of z .



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Table entry for z is the area under the standard normal curve to the left of z .

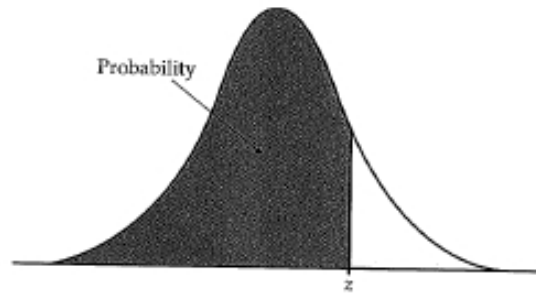


TABLE A										
Standard normal probabilities (continued)										
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998