Statistics 2
First Midterm Exam
Spring 2002

Statistics 2 _
Spring 2002 _
Midterm 1

Printed Name
(Please also p op of each page)

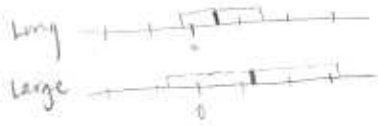
Signature __

Student ID #_ _____

Circle your TA arse Pawel Lasieki Gang Liang
James Van Campen

1	al
20	6

Show your work. No credit will be given for correct answers without justification. A correct answer with no justification or incorrect reasoning will receive no credit.



Rossan Chen

Problem 1. **[6]** In the years 1926-1997, the yearly average increase in value of large-company stocks was 13% and the standard deviation was 20.3%. For long- term corporate bonds, the average increase was 6.1% with an SD of 8.7%. Both histograms are roughly bell-shaped. Is the fraction of years in which large-company stocks lost money larger, about the same as, or smaller than the fraction of years in which long term corporate bonds lost money? Circle one and explain your answer briefly

$\frac{13}{20.3} = 1$ SD

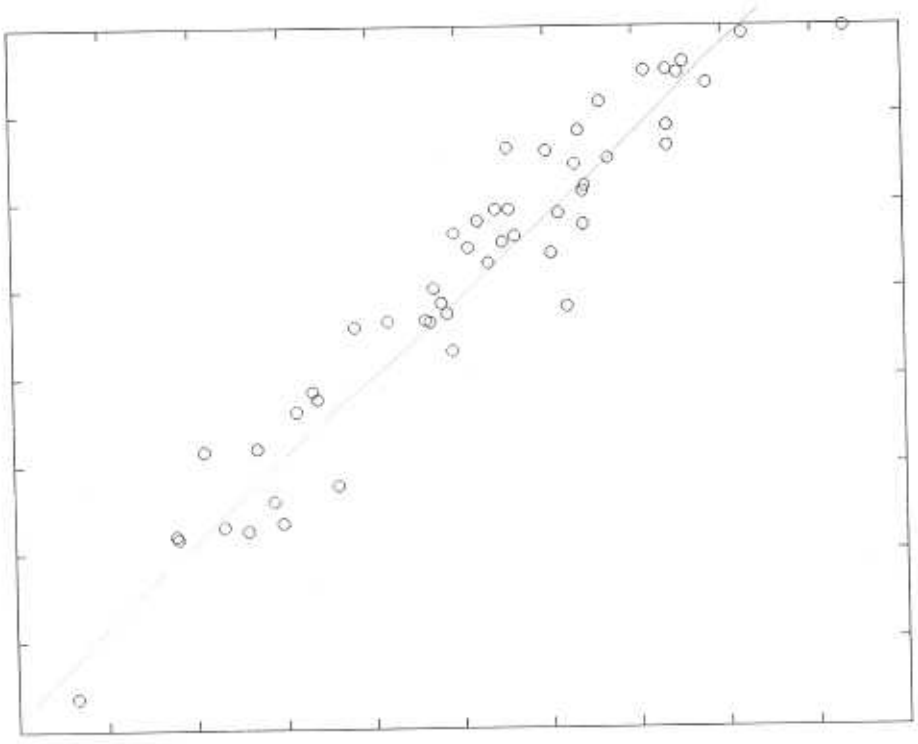
$\frac{7.6}{17.4} = 1$ SD

Same Smaller (Larger)

Large 13%, SD = 20.3%
Long 6.1%, SD = 8.7%

Even though the larger companies made more on average, they also had a higher chance of losing money (as seen in larger SD). Thus, the large companies extended 7.3% into the negative while the long term bonds only extended 2.6% into the negative [see graph in upper left corner]

Problem 2. **[6]** 50 observations are taken on two variables and it is found that they have a correlation coefficient equal to .95. True or false and explain briefly: The scatterplot is roughly similar to the one shown below.



True. All the points fall roughly along a line. Although the points do not fall exactly on the line ($r = 1.0$), they are close enough to warrant a correlation coefficient equal to 0.95.

The graph needs labels, though, to be sure.

Rossan Chen

Problem 3 Read the following and then circle True or False and explain briefly.

---

8 hours of sleep is termed unneeded
Many who get less live longer, study says

Friday, February 15, 2002

Shankar Vedantam, Washington Post

Contrary to popular belief, people who sleep six to seven hours a night live longer, and those who sleep eight hours or more die younger, according to a California study, the largest ever conducted on the subject.

longitudinal

The controversial study, which tracked the sleeping habits of 1.1 million Americans for six years, undermines the advice of many sleep doctors who have long recommended that people get eight or nine hours of sleep every night.

"There's an old idea that people should sleep eight hours a night, which has no more scientific basis than the gold at the end of the rainbow," said Daniel Kripke, a professor of psychiatry at the University of California at San Diego who led the study, published in today's Archives of General Psychiatry.

The study used data from a monumental survey conducted by the American Cancer Society between 1982 and 1988. Women sleeping eight, nine and 10 hours a night had 13 percent, 23 percent and 41 percent higher risk of dying, respectively, than those who slept seven hours, the study found. Men sleeping eight, nine and 10 hours a night had 12 percent, 17 percent and 34 percent greater risk of dying within the period.

(a) [T (F)] [2] The difference between long and short sleepers might be explained by the placebo effect.
There is no placebo effect because all the subjects know how many hours of sleep they are getting. This knowledge does not affect their likelihood of dying

(b) [T (F)] [2] A possible confounding factor is that the participants might not have reported the amount they slept accurately.
Although inaccurate or inconsistent self-reporting may affect the data, it does not help to explain the relationship between the variables (death risk and sleep time)

(c) [T (F)] [2] This is a controlled experiment, but it is not double-blinded
This is an observational study where scientists did not manipulate the subjects' amount of sleep.

(d) [(T) F] [2] A possible confounding factor is that healthier people sleep more.
This statement would actually go against the conclusions of the study, but it is considered a confounding factor because it helps explain the relationship between the variables (death risk and sleep time)

Rossan Chen

Problem 4 [6] This year, a survey was done of men and women who had been divorced and their ages at the time of the divorce were recorded. The results are shown in the figure below; the bars for men are light gray and the bars for women are black.

True or false and explain: Suppose a young woman is married today and is later divorced. She is more likely to be divorced between the ages of 35 and 49 than between the ages of 50 and 69.

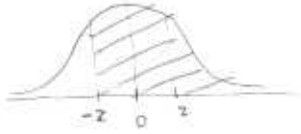
False. This is a cross-sectional, observational study. It is a snap shot in time of different people. Therefore, we cannot conclude upon causation, or apply the data to a person as if it were a longitudinal study (that tracked the same people across time).

6

Problem 5 [6]. As part of a quality control program, objects produced in a factory are weighed and those weighing more than 20 grams are discarded. On the night shift the average weight is 21 grams with an SD equal to 1 and on the day shift the average weight is 21.5 and the SD is 1. Equal numbers of items are produced on both shifts and both histograms are bell shaped. What percentage of the discarded items come from the night shift?

Night: Avg. weight = 21g SD = 1g

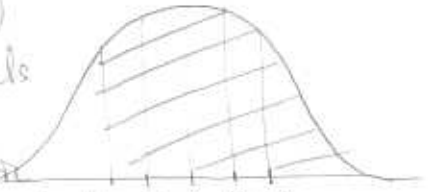
Day: Avg weight = 21.5g SD = 1g



Night: $\frac{21-20g}{1g} = 1\,SD$ If $z = 1.0$, $A = 68.27$

Day: $\frac{21.5-20g}{1g} = 1.5\,SD$ If $z = 1.5$, $A = 86.6$

Need shaded intervals not symmetric intervals



At night, 68% are tossed.
During the day, 87% are tossed.

$\frac{68}{68+87} \times 100\% = 43.8\%$ of discarded items came from night shift

Assume a total of 100 products per shift

Right idea

At night, 84.135% are tossed.
During the day, 93.3% are tossed.

$\frac{84.135}{84.135+93.3} = 47.4\%$ of discarded items came from night shift

Midterm 75 SD=10
Final 70 SD=8
r=0.50

Rossana Chan

Problem 6. Grades on a midterm and a final have a correlation coefficient equal to .50. The average grade on the midterm is 75 and the SD is 10. On the final, the average score is 70 and the SD is 8. The scatterplot is football shaped. A student scores in the 75th percentile on the midterm.

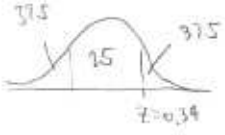
(a) **[4]** What would you predict that student's percentile score to be on the final?

4

75th percentile, z=0.68, 75 + (0.68)10 = 81.2 pts
on midterm

Predicted final score = $70 + \left(\frac{81.2-75}{10}\right) 0.5 \cdot 8 = 72.72$ pts

$\frac{72.72-70}{8} = 0.34$ SD



37.5 + 25 = 62.5

About the 63rd percentile

(b) **[4]** Approximately, what is the chance the student will be above the 90th percentile on the final?

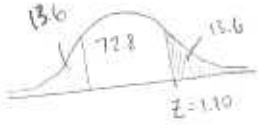
4

90th percentile, z=1.3; 70 + (1.3)(8) = 80.4
on final

Predicted final score for student = 72.72 pts. [see part a]
[new avg.]

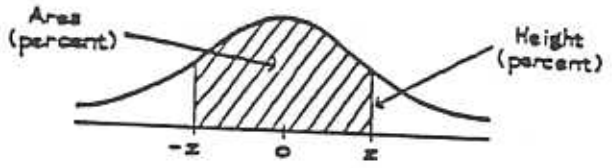
[New SD] = RMS = $\sqrt{1-(.5)^2} \times 8 = 6.93$ pts

$\frac{80.4-72.72}{6.93} = 1.10$ SD



13.6%

Table



A NORMAL TABLE

z	Height	Area	z	Height	Area	z	Height	Area
0.00	39.89	0	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.69	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978
0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	99.9973
1.25	18.26	78.87	2.75	0.91	99.40	4.25	0.005	99.9979
1.30	17.14	80.64	2.80	0.79	99.49	4.30	0.004	99.9983
1.35	16.04	82.30	2.85	0.69	99.56	4.35	0.003	99.9986
1.40	14.97	83.85	2.90	0.60	99.63	4.40	0.002	99.9989
1.45	13.94	85.29	2.95	0.51	99.68	4.45	0.002	99.9991