Your Name: _____

Your Section: _____

Your Score:

Problem	Total	Correct
1	8	
2	27	
3	14	
4	14	
5	14	
6	9	
7	14	
Total	100	

36-201: Introduction to Statistical Reasoning
Final Exam
May 13, 1996

- This exam is closed book, closed notes. You are allowed to use use one 8.5" by 11" sheet of notes (both sides), and a calculator.
- *Show your work* for numerical calculations, and justify your qualitative answers. You will not get full credit unless you do. Also you cannot get any partial credit for a wrong answer if you do not show work. Use the space provided, or the backs of the sheets, to show work and justify answers.
- You do not have to do the problems "in order." Some people like to choose the problems they are going to do best on to work on first, and save the others for later.

*** * * CALCULATORS MAY NOT BE SHARED * * ***

- Graded final exams can be picked up from Dr. Junker (232C Baker Hall) after May 28, 1996.
- You can find out your course grade early by:
  - turning in a stamped self-addressed envelope with your exam today; or
  - bringing a stamped self-addressed envelope to my office, 232C Baker Hall.

1. [8 points; 2 each part]. For each part, mark the single best answer.

(a) In significance testing, a small p-value is bad news for the (mark one box only)

- ☐ Null Hypothesis, $\boldsymbol{H}_0$.
- ☐ Alternative Hypothesis, $\boldsymbol{H}_A$.

(b) A dietician is inspecting the weights of 20 human subjects before beginning an experiment to examine weight gain for people who eat exclusively "fast food". She notices that one subject's weight is recorded as 1180 pounds, and changes the value to 118 pounds after re-weighing that subject. The (mark one box only)

- ☐ mean and standard deviation
- ☐ median and IQR
- ☐ both of the above
- ☐ neither of the above

are likely to be greatly affected by this change.

(c) A simple random sample of 100 CMU students is asked whether they would like more reading days during finals. A 95% percent confidence interval for the proportion that prefer more reading days will be (mark one box only)

- ☐ wider
- ☐ narrower

than a 90% confidence interval for this proportion.

(d) If the survey in part (c) were based on 400 students instead of 100 students, the 95% confidence interval based on 400 students would probably be (mark one box only)

- ☐ wider
- ☐ narrower

than the 95% confidence interval based on 100 students.

2. [27 points] Each situation below describes an experiment or a quasi-experiment. Read each one and answer the accompanying questions. You may use English only if you like, but be clear and specific in your answers.

(a) 50 overweight subjects took part in a study of the effectiveness of weight-loss programs. 25 subjects were randomly assigned to attend a group support program for 10 weeks; the other 25 simply went on about their lives. Private weighings determined each subject's weight at the beginning of the 10 weeks, and again six months after the end of the 10 weeks. For each of the 50 subjects, the change in weight from the beginning of the study to the end of the 6 months was recorded.

(i) [4 points] State an appropriate Null Hypothesis ($\boldsymbol{H}_o$) and Alternative Hypothesis ($\boldsymbol{H}_A$).

$\boldsymbol{H}_o$:_____

$\boldsymbol{H}_A$:_____

(ii) [3 points] For this study, a significance test was performed, with the test statistic (effect size) 3.53, and p-value = 0.00042. Briefly interpret these results.

(iii) [2 points] How should the 50 overweight people have been selected to begin with?

(b) The "Breathalyzer" is used to ascertain the degree of alcohol intoxication from a sample of a person's breath. One of the early uses of the Breathalyzer was in a crackdown on drunken driving undertaken by the British government beginning in September 1967. Police administered the Breathalyzer test to drivers stopped on suspicion, and if it showed intoxication the driver was taken to the police station for more thorough tests. To assess the effects of the crackdown on driving safety, the number of all weekend highway casualties (deaths plus serious injuries) each month were recorded, from January 1966 to September 1968.

(i) [4 points] State an appropriate Null Hypothesis ($\boldsymbol{H}_o$) and Alternative Hypothesis ($\boldsymbol{H}_A$).

$\boldsymbol{H}_o$:_____

$\boldsymbol{H}_A$:_____

(ii) [2 points] Which of the following tools would you use to study this data? (mark one box only)

- ☐ A two-way table
- ☐ A time series plot
- ☐ A scatter plot
- ☐ A histogram

(iii) [3 points] The data show that after the crackdown there was a drop in weekend highway casualties to a level that is about 30% of the pre-crackdown level. Does this show that the Breathalyzer crackdown reduced highway casualties in Great Britain? Explain briefly.

(c) The Cleveland Metropolitan Hospital (Metro) has four similar groups of general internal medicine doctors. From 1976 until 1981 all new patients in internal medicine were assigned randomly to the four groups, to even out the workload. Normally, intravenous (IV) therapy[1] is managed by the regular nurses and physicians in each group. Some physicians thought that adding specialized IV teams to each group would reduce the risk of infections and other complications (like collapsed veins) from IV therapy. So IV teams were added to two of the groups; and the other two groups continued as usual. The doctors kept track of how many new patients undergoing IV therapy in each of the four groups did or did not have infections and other complications.

(i) [4 points] State an appropriate Null Hypothesis ($\boldsymbol{H}_o$) and Alternative Hypothesis ($\boldsymbol{H}_A$).

$\boldsymbol{H}_o$:_____

$\boldsymbol{H}_A$:_____

(ii) [2 points] Which of the following tools would you use to study this data? (mark one box only)

- ☐ Confidence intervals for means
- ☐ Chi-squared test
- ☐ Correlation
- ☐ Can't tell from the information given

(iii) [3 points] Why was it a good idea, for this experiment, that new internal medicine patients were randomly assigned to each group.

[1] Feeding patients fluids by tube into an arm vein.

3. [14 points] The Honolulu Heart Program has conducted a variety of studies of factors that may cause or prevent strokes[2], since the late 1960's. One study enrolled 3,150 men of Japanese ancestry living on the Hawaiian island of Oahu, and watched them for 22 years. The men ranged from 55 to 68 years of age when the study began in 1965. One of the factors that was watched was the amount of milk each man drank per day. At the end of the study, the number of men who had suffered a thromboembolic stroke [the most common type of stroke] were counted. Table 1 shows how many milk drinkers had strokes, how many did not, etc.

Observed Counts:

	Stroke?		
Milk Consumption	Yes	No	Total
$\geq$ 2 glasses per day	18	463	481
$<$ 2 glasses per day	211	2458	2669
Total	229	2921	3150

Table 1: From an article by Terry Wallace, Associated Press, published in the *Pittsburgh Post-Gazette,* May 7, 1996.

(a) [3 points] This study is a/an (mark one box only)

- ☐ Experiment.
- ☐ Observational Study.
- ☐ Survey.
- ☐ None of the above, or can't say for sure.

(b) [4 points] Fill in the table below with row percents, using the observed counts in Table 1.

Row Percents:

	Stroke?		
Milk Consumption	Yes	No	Total
$\geq$ 2 glasses per day			100%
$<$ 2 glasses per day			100%

[2] In medicine, a *stroke* is a sudden loss of muscular control, with diminution or loss of sensation and consciousness, resulting from rupture or blocking of a blood vessel in the brain.

(c) [3 points] Based on the row percents you calculated in part (b), is there *positive*, *negative* or *no* association between milk consumption and suffering a stroke? Use specific numbers from part (b) in your answer.

(d) [4 points] This study was published in the May 1996 issue of the medical journal *Stroke*. In the journal article, the authors do *not* conclude that drinking milk causes a lower risk of stroke. Name a very plausible lurking variable that could interfere with a causal connection between drinking milk and reducing strokes, and draw a diagram to show whether your lurking variable influences stroke only (confounding) or both milk drinking and stroke (common response).

4. [14 points] The presence of sulfur compounds in wine is a defect that causes wine lovers to describe the odor of the wine as "cabbage" or "onion". Wine producers would like to know the lowest concentration of these compounds that can be reliably detected by trained wine tasters; this is called the "odor threshhold". In a study to determine the odor threshhold, a trained wine judge is presented with three glasses of wine. One is identified as containing just the wine and no sulphur compounds, for reference. The remaining two glasses are not identified: one contains just the wine, and in the other the wine has been spiked with a fixed concentration of the sulfur compound. The judge must determine by taste which glass was spiked. This setup is repeated five times for each judge, with increasing concentrations of sulphur determined in advance by the wine producer.

20 judges took part in one such study to determine the odor threshhold for the compound ethanethiol. Table 2 below gives the percent of judges who correctly identified the spiked wine at each concentration level, in micrograms per liter (the actual concentration levels are not given; instead the logarithms of the concentration levels are given). Figure 1 on page 8 shows an analysis of the data using *Data Desk*.

Log-concentration	0.3	0.5	0.6	0.7	0.8
% correct	55	65	75	70	85

Table 2: Percent of judges correctly identifying spiked wine in wine-tasting experiment. Based on O. J. Goniak and A. C. Noble (1987), "Sensory study of selected volatile sulphur compounds in white wine," *American Journal of Enology and Viticulture, 38,* pp. 223–227.

(a) [3 points] This study is a/an (mark one box only)

- ☐ Experiment.
- ☐ Survey.
- ☐ Census.
- ☐ None of the above, or can't say for sure.

(b) [3 points] In this study, (mark one box only)

- ☐ concentration of ethanethiol is the *explanatory variable* and percent correct is the *response variable.*
- ☐ percent correct is the *explanatory variable* and concentration of ethanethiol is the *response variable.*
- ☐ it doesn't make sense to identify explanatory and response variables.

[continued on page 9.]

Figure 1: *Data Desk* analysis of the wine-tasting data for problem 4.

(c) [4 points] In the *Data Desk* output in Figure 1 (p. 8), the data from Table 2 has been plotted in a scatter plot and the regression line has been plotted. A correlation is also given. Describe the plot and the correlation. Comment on how well a straight line fits the data, and on how much and what kind of association you see in the data.

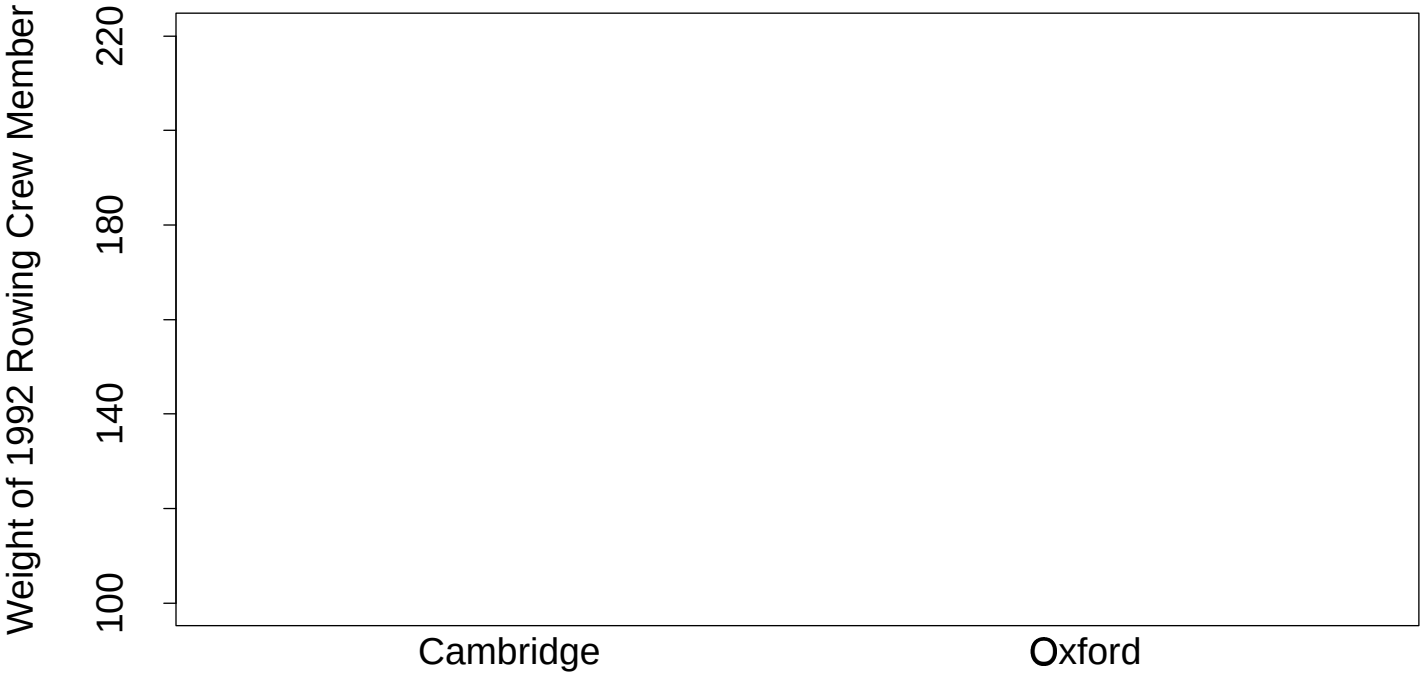
(d) [4 points] The wine producer is interested in determining the log-concentration of ethanethiol that is identified correctly 75% of the time (this is the "odor threshhold"). Use the regression line plotted in the *Data Desk* output in Figure 1 to find this value (you may wish to sketch something on Figure 1 to help explain your answer).

5. [14 points] In a story dated March 31, 1992, the British newspaper *The Independent* compared the rowing crews competing in the Oxford and Cambridge boat race that year. Below is a table of the weights of each crew member in pounds, which crew that person is on (Oxford or Cambridge) and whether that person is a rower (Rower=1) or a coxswain (Rower=0) [the coxswain is the person who steers the boat and has charge of its crew]. Some output from a *Data Desk* analysis of this data is given in Figure 2 on p. 12.

Weight	*School*	*Rower?*
188.5	Cambridge	1
183	Cambridge	1
194.5	Cambridge	1
185	Cambridge	1
214	Cambridge	1
203.5	Cambridge	1
186	Cambridge	1
178.5	Cambridge	1
109	Cambridge	0

Weight	*School*	*Rower?*
186	Oxford	1
184.5	Oxford	1
204	Oxford	1
184.5	Oxford	1
195.5	Oxford	1
202.5	Oxford	1
174	Oxford	1
183	Oxford	1
109.5	Oxford	0

(a) [4 points] In the space provided, make side-by-side boxplots of the Cambridge and Oxford team weights. Be sure to show outliers according to the 1.5∗IQR rule, etc. *Summary statistics given in Figure 2 on p. 12 will be very helpful in constructing the boxplots!*



(b) [3 points] List all the outliers you found in part (a) here, and give a possible reason why each was an outlier. If you can't think of a reason, just put "no obvious reason".

Value	School	Possible Reason

(c) [4 points] Consider the outliers you listed in part (b):

- If the outlier(s) for Cambridge were left out, the mean weight of the remaining Cambridge crew members would be (mark one box only)
  - ☐ about the same
  - ☐ somewhat higher
  - ☐ somewhat lower

  than the current mean of 182.4 pounds.
- If the outlier(s) for Oxford were left out, the mean weight of the remaining Oxford crew members would be (mark one box only)
  - ☐ about the same
  - ☐ somewhat higher
  - ☐ somewhat lower

  than the current mean of 180.4 pounds.

(d) [3 points] In Figure 2 there is some *Data Desk* output showing a confidence interval for the difference in mean weight of rowing crew members from each school, and a significance test for

$\boldsymbol{H}_0$**:** the mean weight of rowing crew members at Cambridge and Oxford are the same; vs.
$\boldsymbol{H}_A$**:** the two means are different.

Can you use the confidence interval and hypothesis test to draw conclusions about the populations of all rowing crew members at each school? Explain.

Figure 2: *Data Desk* output for the rowing crew problem (problem 5).

Figure 3: "Ask Marilyn" column for problem 6.

6. [9 points] Figure 3 above shows a question that was published in the newspaper column "Ask Marylin", by Marilyn Vos Savant, in the *Parade* Sunday Magazine of the *Pittsburgh Post-Gazette*, April 26, 1996. (Marilyn's response has been omitted.)

(a) [5 points] On the next page are tables of observed counts and of row percents for this problem. Some of the observed counts are still missing. Fill in the missing numbers.

Observed Counts

Applicant:	White Collar Job Hired? Yes	White Collar Job Hired? No	White Collar Job Total	Blue Collar Job Hired? Yes	Blue Collar Job Hired? No	Blue Collar Job Total	Combined Hired? Yes	Combined Hired? No	Combined Total
Male			200	300	100	400	330	270	600
Female	40	160	200	85	15	100			
Total	70	330	400	385	115	500	455	445	900

Row Percents

Applicant:	White Collar Job Hired? Yes	White Collar Job Hired? No	White Collar Job Total	Blue Collar Job Hired? Yes	Blue Collar Job Hired? No	Blue Collar Job Total	Combined Hired? Yes	Combined Hired? No	Combined Total
Male	15%	85%	100%	75%	25%	100%	55.00%	45.00%	100%
Female	20%	80%	100%	85%	15%	100%	41.67%	58.33%	100%

(b) [4 points] This is an example of Simpson's Paradox. Write a response from Marilyn, using ideas and concepts we have learned in class, and referring to specific numbers in the tables above. Remember, this is in *Parade* magazine, so your explanation has to be understandable to someone with a high school education or less!

7. [10 points] In Figure 4 on p. 16 is a summary of a recent New York Times / CBS News poll comparing voters' perceptions of Bill Clinton and Bob Dole, the leading candidates for President in the November elections. The accompanying *New York Times* article states [paraphrasing slightly],

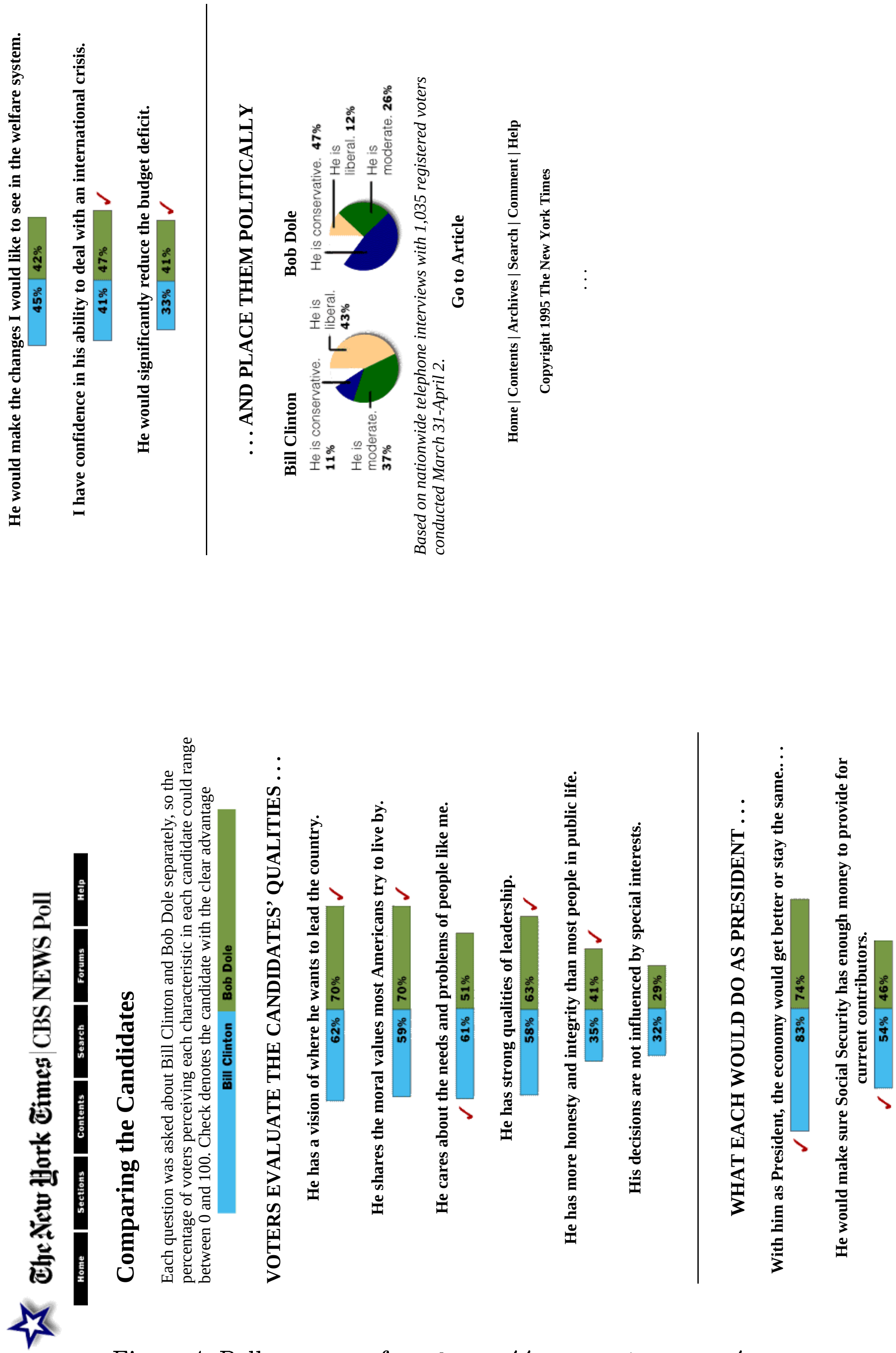
> This survey is based on telephone interviews conducted March 31 to April 2 with 1,035 adults throughout the United States who said they were registered to vote. The sample of telephone exchanges called was randomly selected by a computer from a complete list of more than 36,000 active residential exchanges across the country. Within each exchange, random digits were added to form a complete telephone number, thus permitting access to both listed and unlisted numbers. Within each household, one adult was designated by a random procedure to be the respondent for the survey. In theory, in 19 cases out of 20 the results based on such samples will differ by no more than three percentage points in either direction from what would have been obtained by seeking out all American adults. For smaller subgroups the potential sampling error is larger.

Please answer the following four questions based on the figure and the information given above.

(a) [7 points, one per blank] Fill in the blanks (if you can't tell, just write "can't tell"):

- What was the population? (What/who were the units or subjects?) _____.
- How was the sample selected? (SRS, self-selected, stratified, multistage, or ...?) _____.
- How were the sample subjects contacted? (Telephone, US mail, email, ...?) _____.
- What was the sample size? _____.
- What was the overall margin of error? _____.
- When was the survey conducted? _____.
- Can you find out the exact survey questions asked? _____.

[Continued on page 17]



The New York Times | CBS NEWS Poll

Home | Sections | Contents | Search | Forums | Help

Comparing the Candidates

Each question was asked about Bill Clinton and Bob Dole separately, so the percentage of voters perceiving each characteristic in each candidate could range between 0 and 100. Check denotes the candidate with the clear advantage

Bill Clinton	Bob Dole

VOTERS EVALUATE THE CANDIDATES' QUALITIES . . .

Statement	Bill Clinton	Bob Dole
He has a vision of where he wants to lead the country.	62%	70% ✓
He shares the moral values most Americans try to live by.	59%	70% ✓
He cares about the needs and problems of people like me.	✓ 61%	51%
He has strong qualities of leadership.	58%	63% ✓
He has more honesty and integrity than most people in public life.	35%	41% ✓
His decisions are not influenced by special interests.	32%	29%

WHAT EACH WOULD DO AS PRESIDENT . . .

Statement	Bill Clinton	Bob Dole
With him as President, the economy would get better or stay the same.. . .	✓ 83%	74%
He would make sure Social Security has enough money to provide for current contributors.	✓ 54%	46%
He would make the changes I would like to see in the welfare system.	45%	42%
I have confidence in his ability to deal with an international crisis.	41%	47% ✓
He would significantly reduce the budget deficit.	33%	41% ✓

. . . AND PLACE THEM POLITICALLY

Bill Clinton

He is conservative. 11%
He is liberal. 43%
He is moderate. 37%

Bob Dole

He is conservative. 47%
He is liberal. 12%
He is moderate. 26%

Based on nationwide telephone interviews with 1,035 registered voters conducted March 31-April 2.

Go to Article

Home | Contents | Archives | Search | Comment | Help

Copyright 1995 The New York Times

. . .

Figure 4: Poll summary from `http://www.nytimes.com/`.

(b) [4 points] Use the margin of error you wrote down in part (a) to calculate the following confidence intervals (these two questions are at the top of the second column in the figure).

- For the question *"He would make the changes I would like to see in the welfare system"*,

  The proportion of voters who think this about Clinton

  is between _____% and _____%, with 95% confidence.

  The proportion of voters who think this about Dole

  is between _____% and _____%, with 95% confidence.

- For the question *"I have confidence in his ability to deal with an international crisis"*,

  The proportion of voters who think this about Clinton

  is between _____% and _____%, with 95% confidence.

  The proportion of voters who think this about Dole

  is between _____% and _____%, with 95% confidence.

(c) [3 points] In the figure, under the *"international crisis"* question there is a check ("$\surd$") near Dole's percentage but not Clinton's; and there is no check at all under the *"welfare"* question. Use your answers to part (c) to explain what the check means.