

Name _____

Student # _____

STAT 101: Final Examination

Ques	Mark	Out of
1		5
2		5
3		4
4		4
5		1
6		1
7		5
8		5
9		5
10		5
11		3
12		2
13		5

Instructions:

1. This is an open book exam.
2. You may use a calculator (with no wireless communications ability).
3. You may *not* use a phone as a calculator or in any other way.
4. Sometimes, to do the problem, you will need to make assumptions. You should be clear and explicit about what assumptions you need to make the technique you are using reasonable. Marks will be deducted for failing to make such assumptions clear. The problems where you need to say something about assumptions are not labelled.
5. The exam is out of 50.
6. There are 15 pages including this one.

Questions number 1 to 7 refer to the following experiment. The data are from the Data Stories Library hosted at Statlib.

Two identical footballs, one air-filled and one helium-filled, were used outdoors on a windless day at The Ohio State University's athletic complex. Each football was kicked 39 times and the two footballs were alternated with each kick. The experimenter recorded the distance travelled (in yards) by each ball. Here are the data, some summary statistics, histograms for the variables and some scatterplots.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13
Air	25	23	18	16	35	15	26	24	24	28	25	19	27
Helium	25	16	25	14	23	29	25	26	22	26	12	28	28
Difference	0	7	-7	2	12	-14	1	-2	2	2	13	-9	-1
Trial #	14	15	16	17	18	19	20	21	22	23	24	25	26
Air	25	34	26	20	22	33	29	31	27	22	29	28	29
Helium	31	22	29	23	26	35	24	31	34	39	32	14	28
Difference	-6	12	-3	-3	-4	-2	5	0	-7	-17	-3	14	1
Trial #	27	28	29	30	31	32	33	34	35	36	37	38	39
Air	22	31	25	20	27	26	28	32	28	25	31	28	28
Helium	30	27	33	11	26	32	30	29	30	29	29	30	26
Difference	-8	4	-8	9	1	-6	-2	3	-2	-4	2	-2	2

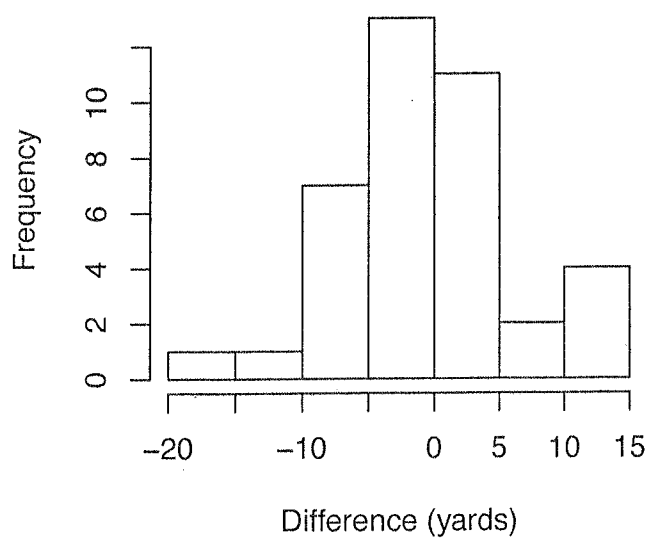
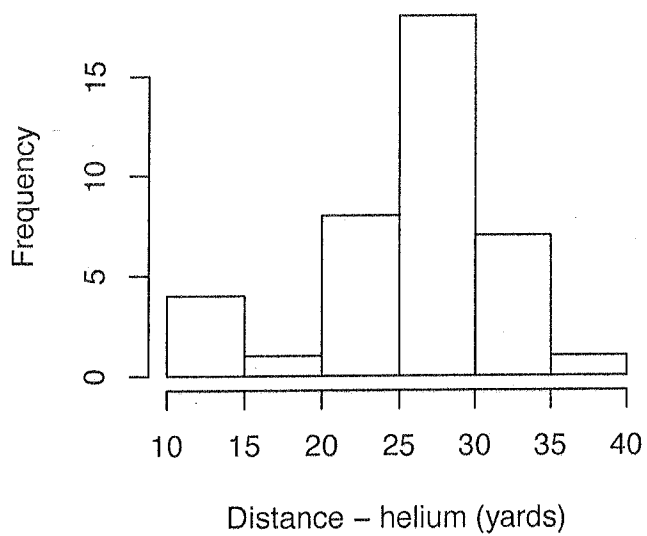
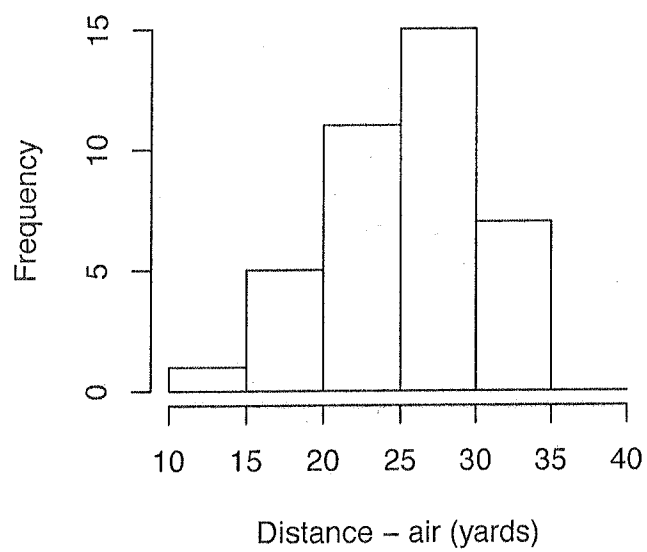
Here are some summary statistics:

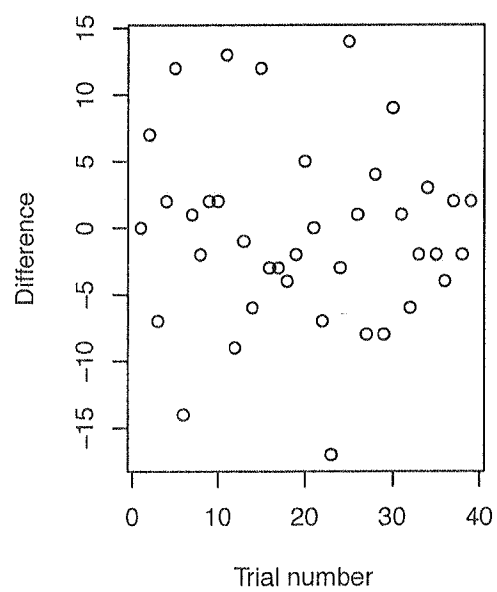
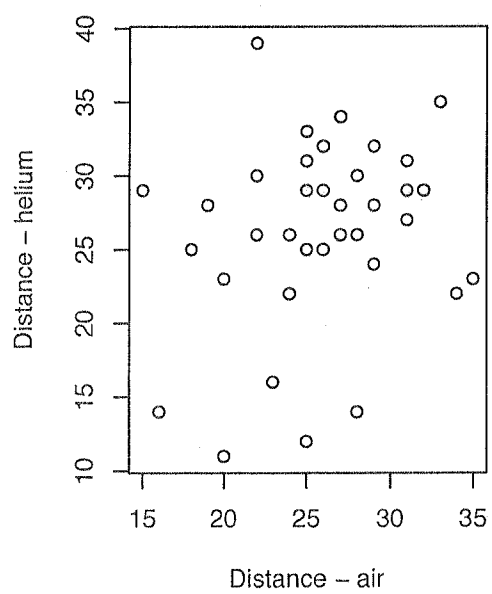
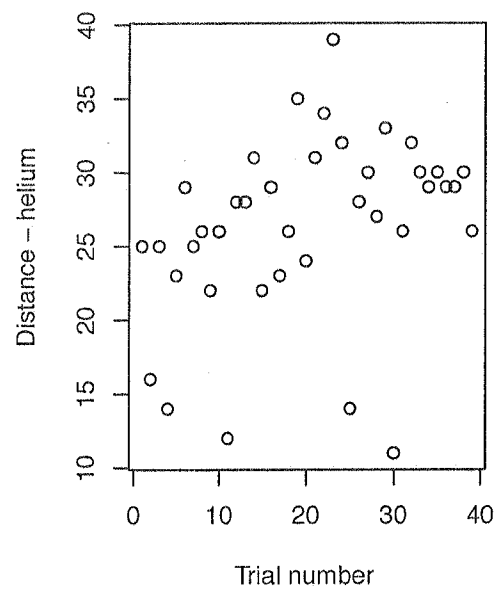
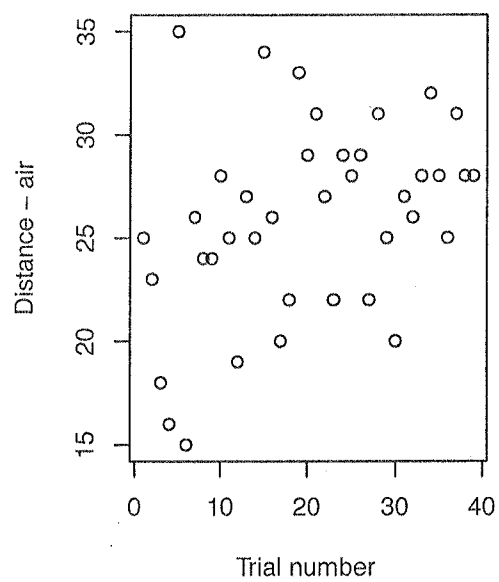
	Trial	Air	Helium	Difference
Mean	20.0	25.92	26.38	-0.4615
SD	11.40	4.69	6.21	6.87
Median	20	26	28	-1

The correlation between distance in air and distance with helium is 0.23.

The correlation between trial number and distance with air is 0.36.

The correlation between trial number and distance with helium is 0.34.





1. Does using helium affect the distance the ball travels? [5 marks]

2. Give a 95% confidence interval for the distance a helium filled ball travels. [5 marks]

3. Make a normal approximation to estimate the fraction of trials for which the air ball travels more than 5 yards further than the helium ball. [4 marks]

4. Use regression to predict the distance a helium filled ball will travel on a trial where the air filled ball travels 45 yards. [4 marks]

5. Suggest one clear danger in using this method of prediction. [1 mark]
6. Explain why a paired comparisons (matched pairs) design was useful in this case. Your answer will refer to the data and/or the plots. [1 mark]
7. In 14 of the 39 trials the kick with the air filled ball is longer than the kick with the helium filled ball. Give a 95% confidence interval for the probability that the air filled ball will be kicked farther than the helium filled ball on a given trial. [5 marks]

8. In the General Social Survey a sample of adults were asked whether or not they supported the statement: "It is right to use animals for medical testing of it might save human lives." The responses, broken down by sex, are:

	Men	Women
Supportive	346	306
Sample Size	516	636

Find an 80% confidence interval for the difference between the population proportions of men and women who support this statement. [5 marks]

9. The following story is taken from the data story library at lib.stat.cmu.edu.
 “A manufacturer was considering marketing crackers high in a certain kind of edible fiber as a dieting aid. Dieters would consume some crackers before a meal, filling their stomachs so that they would feel less hungry and eat less. A laboratory studied whether people would in fact eat less in this way. Overweight female subjects ate crackers with different types of fiber (bran fiber, gum fiber, both, and a control cracker) and were then allowed to eat as much as they wished from a prepared menu. The amount of food they consumed and their weight were monitored, along with any side effects they reported. Unfortunately, some subjects developed uncomfortable bloating and gastric upset from some of the fiber crackers. A contingency table of ‘Cracker’ versus ‘Bloat’ shows the relationship between the four different types of cracker and the four levels of severity of bloating as reported by the subjects.”

The data are:

	Bloating				Total
	High	Low	Medium	None	
Bran	0	4	1	7	12
Combo	2	5	3	2	12
Control	0	4	2	6	12
Gum	5	2	3	2	12
Total	7	15	9	17	48

Carry out a χ^2 test to see if there is a relationship between type of cracker and the extent of bloating symptoms.

As part of your answer you should complete the following tables. [5 marks]

Expected Cell Counts

	Bloating			
	High	Low	Medium	None
Bran	1.75	3.75	2.25	4.25
Combo				4.25
Control				
Gum				

Components of X^2

	Bloating			
	High	Low	Medium	None
Bran	1.75	0.02	0.69	1.78
Combo	0.04		0.42	
Control				
Gum	6.04	0.82	0.25	1.19

10. A sample of 20 men is drawn at random from a large population of men. An independent sample of 20 women is also drawn. The individuals' brain sizes (in unknown units) were measured along with their IQs. The data are below.

	Men		Women	
	IQ	Brain Size	IQ	Brain Size
	140	1001	133	817
	139	1038	137	952
	133	965	99	929
	89	909	138	991
	133	955	92	854
	141	1080	132	834
	135	924	140	856
	100	945	96	879
	80	889	83	865
	83	892	132	852
	97	906	101	808
	139	955	135	791
	141	935	91	832
	103	1062	85	799
	144	950	77	794
	103	998	130	867
	90	880	133	858
	140	950	83	834
	81	930	133	948
	89	936	88	894
Mean	115.0	954.9	111.9	862.7
SD	25.0	55.9	23.7	55.9

Do the men have larger brain sizes than the women in these populations? [5 marks]

11. On the next page you will see output from JMP in which I regressed IQ on Brain Size for the data set in the previous question putting both men and women together. Give a 90% confidence interval for the slope of the regression line of IQ on brain size in this population. [3 marks]

JMP output for IQ and brain size question

Brains: Fit Least Squares

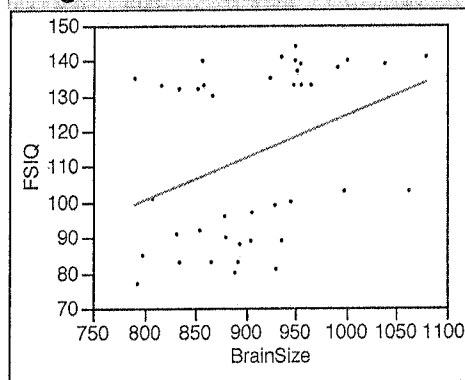
Page 1 of 1

Response FSIQ

Whole Model

BrainSize

Regression Plot



Summary of Fit

RSquare	0.127907
RSquare Adj	0.104957
Root Mean Square Error	22.78325
Mean of Response	113.45
Observations (or Sum Wgts)	40

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	5.1677034	46.00819	0.11	0.9112
BrainSize	0.1191546	0.050472	2.36	0.0235*

12. I have a box with 3 red and 2 green billiard balls in it. I have a bag with 1 red and 3 green balls. I pick one ball from the box without looking and 1 ball from the bag without looking. What is the chance that I get two red balls?
[2 marks]

13. Each time I give STAT 101 or 201 I survey the class in the first lecture. In 2004 I had 154 respondents of whom 100 were female. This year I had 66 of whom 39 were female. Assume that we can treat the respondents each year as a sample of all SFU undergrads who would respond to such a survey if asked. Has the proportion of such willing respondents who are female changed between years?
[5 marks]

