

Solutions

STAT 651 Midterm Test (50 minutes) - 1:50pm - 2:40pm November, 2010

NAME:

Total number of Marks: /25

There are 6 questions in this paper, do not be deterred, they are all straightforward. Read each question carefully. There are questions on both side of the page. The number of marks for each question are given in brackets. Be smart about how you answer. If you can't answer one question move on the to next and return to the questions you could not do after answering all the other questions!

Rubric: This exam is an open book exam you can use all written materials that you want, normal tables and a calculator.

Write your solutions in the question paper.

GOOD LUCK!

(1a) Suppose you want to test the equality of two population means against the alternative that the two population means are different (we will assume that both populations have the same variance). You do not observe the whole population, but you do observe samples of size m from population 1 and n from population 2. Given the information below, which sample sizes should I choose.

(A) The population variances is 1, $m = 50$ and $n = 50$.

(B) The population variances is 0.5 $m = 40$ and $n = 40$

(C) The population variance is 0.5 $m = 40$ and $n = 60$.

(D) The population variance is 1 $m = 45$ and $n = 55$.

$$(A) \text{ s.e} = \sqrt{\frac{1}{50} + \frac{1}{50}} \quad (B) \text{ s.e} = \sqrt{0.5 \left(\frac{1}{40} + \frac{1}{40} \right)}$$

$$(C) \text{ s.e} = \sqrt{0.5 \left(\frac{1}{40} + \frac{1}{60} \right)} \quad (D) \text{ s.e} = \sqrt{\frac{1}{45} + \frac{1}{55}}$$

Smallest s.e is (C)

[2]

(b) Xuan draws 300 samples, each sample is of size 30. For each sample, Xuan constructs a 95% CI, on average how many of these confidence intervals will contain the true mean?

[1]

$$300 \times 0.95 = 285$$

(c) Jake reads about an opinion poll, that says the proportion of Americans who were optimistic about their future is $[0.58 \pm 0.1]$ with 95% confidence.

Jake asks you what they mean by 'the proportion of Americans who were optimistic about their future is $[0.58 \pm 0.1]$ with 95% confidence', and what would happen to this interval if the number of people sampled in the opinion poll increased. Answer his questions in four lines or less.

The estimator of the proportion of Americans who were [2]

optimistic is based on a sample. The estimator is 0.58.

But with 95% confidence we believe the true proportion lies in the interval $[0.48, 0.68]$.

If we increase the size (number of people interviewed) of the sample, the length of the CI decreases.

- (2) Two brothers, Daniel and Joseph are having a debate about Daniel's frog. Daniel claims that his frog has psychic powers, because his frog has correctly predicted the outcome of the past four baseball matches (assume here that each match has only two teams, team A and team B, taking part, and the outcome can only be team A or team B winning - a draw is not allowed). Joseph disputes his brother's claim.

Help Joseph win the argument. By stating the null and alternative hypothesis for this scenario, use the ideas behind statistical testing to show that based on the frog's predictions, there is no evidence that Daniel's frog is psychic.

H_0 : Frog is not psychic. H_A : Frog is psychic.

If the frog is not psychic, the probability it will predict the outcome of any one given match correctly is $\frac{1}{2}$.

Thus the probability he will predict 4 matches correctly (just by chance) is $(\frac{1}{2})^4 = \frac{1}{16} = 6.25\%$. As this probability is not too small, (6.25% is quite large), there is no evidence to back the alternative (that the frog is psychic). [3]

- (3) Suppose I do a hypothesis test and I am unable to reject the null at the 5% level. Which statement is correct? [2]

(A) A type I error could have been committed, the p -value is less than 5% and I know the type I error.

→ (B) A type II error could have been committed, and the p -value is greater than 5%. ←

(C) A type II was committed, and the p -value is greater than 5%.

(D) A type II error could have been committed, the p -value is greater than 5% and I know the type II error.

- (4) Let μ denote the population mean and \bar{X} denote the sample mean. Suppose we observe a sample of 25 individuals, and the sample mean is $\bar{X} = 3$. Bill wants to test the research hypothesis that the mean is greater than two. He states the hypothesis as $H_0: \bar{X} \leq 2$ against the alternative $H_A: \bar{X} > 2$. Did he write his hypothesis correctly? Give a reason for your answer. [1]

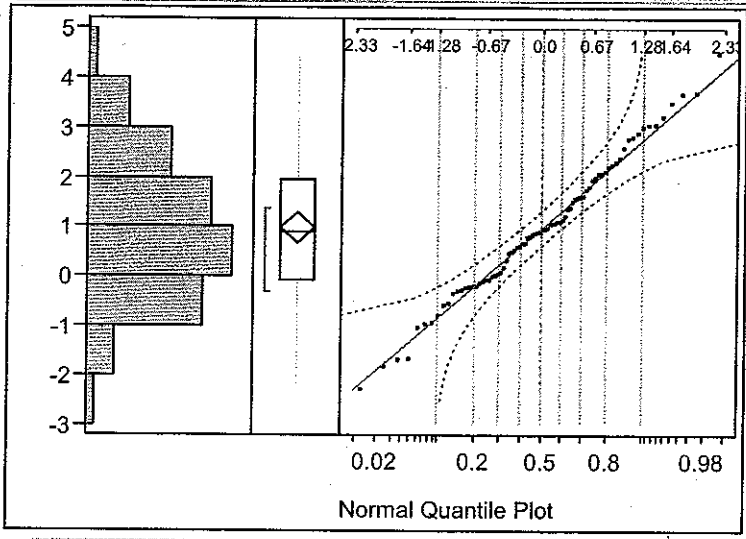
No. Bill states the hypothesis in terms of the sample mean \bar{X} , which is known ($\bar{X} = 3$). We are interested in the population mean (which is unknown). Based on $\bar{X} = 3$ we want to test

$H_0: \mu \leq 2$ against $H_A: \mu > 2$.

Distributions

Question 5

Column 2



→ Data look quite normal and sample size $n=80$ is large. For this reason by CLT \bar{X} is close to normal.

Moments

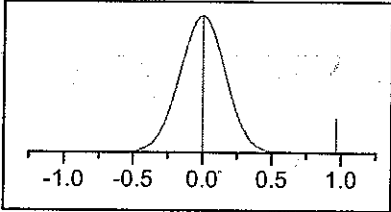
Mean	0.965875
Std Dev	1.3867024
Std Err Mean	0.155038
Upper 95% Mean	1.2744705
Lower 95% Mean	0.6572795
N	80

→ $s = 1.39$ estimator of standard deviation σ .
 → $s.e. = 0.155$.
 → sample size = 80. This is large enough for us to use the normal distribution rather than the t-distribution with 79 degrees of freedom.

Test Mean=value

Hypothesized Value	0
Actual Estimate	0.96588
DF	79
Std Dev	1.3867
t Test	
Test Statistic	6.2299
Prob > t	<.0001*
Prob > t	<.0001*
Prob < t	1.0000

→ Test is at zero not 0.9.



[Faint handwritten notes and bleed-through from the reverse side of the page are visible in this area.]

- (6) Roopa is doing research on the amount of iron gained or lost after going on two different types of diets. She took a sample of 20 volunteers and randomly allocated them into two groups (ten people in each group). Group 1 she put on a high vitamin C diet, the amount of iron gained or lost is given in the first row below (if the number is positive this is a gain, if the number is negative it is a loss). Group 0 she put on a high calcium diet, the iron gained or lost is given in the second row of the table below.

group										
1	0.51	2.75	0.79	4.41	-1.23	1.06	1.98	2.32	1.59	-18.41
0	-0.18	0.92	-0.25	1.56	-0.38	-0.21	-0.62	-1.68	-3.15	-0.33

- (i) Roopa's research hypothesis is that there is a difference between the average amount of iron gained for the two diets. Let μ_1 denote the mean amount of iron gained on the vitamin C diet and μ_0 denote the mean amount of iron gained on the high calcium diet.

State Roopa's null and alternative hypothesis?

$$H_0: \mu_1 - \mu_0 = 0$$

$$H_A: \mu_1 - \mu_0 \neq 0$$

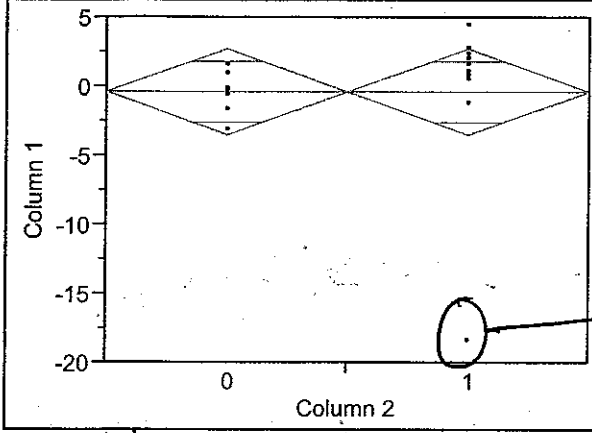
There is a difference in an average ^{of} iron absorption between the [1] diets.

- (ii) State precisely the statistical test(s) that you would use to test this hypothesis. Give a reason for your answer, and the assumptions required.

There is independence between and within the samples. Hence an independent 2-sample test makes sense.

The sample sizes $n, m = 10$ are small. Moreover [2] looking at the data and the plot overlay suggests there are outliers, which would influence the result of a independent sample t-test. Therefore it would appear a Wilcoxon sum rank test is the most appropriate test for this data set.

Oneway Analysis of Column 1 By Column 2



Question 6
part (iii), (iv).

For part (iii) and (iv)
- overleaf

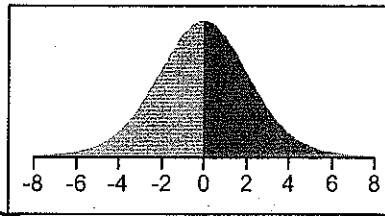
→ Huge outlier, which will drag down the average of Group 1.

Oneway Anova

t Test

1-0
Assuming equal variances

Difference	0.0090	t Ratio	0.004299
Std Err Dif	2.0935	DF	18
Upper CL Dif	4.4072	Prob > t	0.9966
Lower CL Dif	-4.3892	Prob > t	0.4983
Confidence	0.95	Prob < t	0.5017



p-value for the test is $\frac{0.9966}{2} \approx 0.5$

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Score Mean	(Mean - Mean0)/Std0
0	10	78.000	7.8000	-2.003
1	10	132.000	13.2000	2.003

2-Sample Test, Normal Approximation

S	Z	Prob > Z
132	2.00321	0.0452

1-way Test, ChiSquare Approximation

ChiSquare	DF	Prob > ChiSq
4.1657	1	0.0413*

→ Since it is a two-sided test and $n = m = 10$, we can choose either value to do the test. The non-reject region is

Approximate p-value is

$\frac{0.0452}{2} \approx 0.0226$

Since $0.0226 < 0.025$, there is evidence to reject the null.

(iii) Roopa put the data in JMP, and runs some tests. Her JMP output is given on the adjacent page. The tests are done at the 5% level.

What are the results of her tests, when she uses the t-test and the Wilcoxon test?

Independent sample t-test: p-value $\approx 0.5 > 0.025$, hence no evidence to reject the null.

Wilcoxon sum rank test: $T = 132$. Use table 5 for $n = m = 10$.

The non-rejection region is $[79, 131]$. Since $132 > 131$, hence evidence to reject null. [2]

(iv) Do the two tests give different results? If this is the case, by looking at the data and the output, explain why this may be the case. By using the data and the output, which test gives the most reliable answer (give a reason for your answer).

The conclusions of the Independent sample t-test and the Wilcoxon test are different. This is

probably because group 1 contains a huge outlier (-17.41) , which drags down the mean of

Group 1. Here the difference in sample means is $\bar{X} - \bar{Y} = 0.009$, which is very small compared to the s.e

s.e = 2.09. This is the reason we cannot reject the null for the independent sample t-test.

In contrast, contrast, the Wilcoxon test only gives a large rank to -17.41 , but not the magnitude. Hence ~~less~~^{less} weight is given to this observation, it exerts less influence in the test.

Since the data is highly non-normal, with outliers and sample size is small, it appears that the Wilcoxon test is more reliable.