

ADM 2304
Fall 2013
Solutions and Marking Guide to Assignment 1.

Note that Minitab output does not substitute for the different elements of a hypothesis test—it should be viewed strictly as a calculator. The hypotheses, decision and conclusion must be written or typed and show separately from any computer output for marks. No marks are given for a solution that only provides Minitab output.

Problem 1. [9 marks]

(a) Using Minitab, we can calculate the population mean of the MedInc values as 89368.

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
MedInc	1241	0	89368	918	32337	15947	67725	89103	109436
AvgInc	1241	0	100690	1106	38973	18279	75632	97312	120212

If the mean of the AvgInc values is given as 100690, we will accept this as well.

1 mark for mean of \$89368 or \$100690. Give full marks if it appears the manipulations are based on AvgInc instead of MedInc.

(b) *3 marks for evidence of 20 CIs.*

(c) *1 mark for any graph of data*
1 mark for showing the manual computation of the CI
2 marks for any reasonable comment re data distribution— that the small sample can or cannot be assumed to come from a normal distribution or that the large sample came from a population that is not extremely skewed.

(d) *1 mark for count*

Problem 2. [11 marks]

(a) $H_0: p = .3462$, $H_a: p < .3462$
 $n = 1000$ with $\hat{p} = .304$
 $z = (.304 - .3462) / \sqrt{.3462 * .6538 / 1000} = -2.8$
 At the .01 significance level, we reject H_0 since $-2.8 < -2.32$
 Conclude the Conservative vote has dropped by more than 5% from their popular vote in 2011

Test of $p = 0.3462$ vs $p < 0.3462$

Sample	X	N	Sample p	95% Upper Bound	Z-Value	P-Value
1	304	1000	0.304000	0.327926	-2.80	0.003

4 marks:

1 for hypotheses (deduct .5 for 2-sided alternative)

1 for z-statistic or CI or p-value (deduct .5 if z is based on $SE = \sqrt{.304 * .696 / 1000}$ instead of SD based on $p = .3462$.)

1 for decision to reject the null H.

1 for conclusion

(b) [2 marks]

Using p estimated at .304,

we have $n = pq (z / M)^2 = .304 * .696 * (2.576 / 0.01)^2 = 14041$.

Using $p = 0.5$ ($n = 16590$), $p = 0.3962$ ($n = 15875$) or $p = 0.3462$ ($n = 15020$) would not be reasonable at this time, since no federal political party would ever get 50% of the popular vote and we have just rejected the null hypothesis that $p \geq 0.3462$. However, give 1 mark for the calculations in these cases.

(c)

$H_0: p = .3962$, $H_a: p < .3962$

$n = 17$ is not sufficient to assume a normal distribution for p -hat since $np = 17 * .3962 = 6.7$ is not greater or equal to 10 or since $X = 3$ is not greater or equal to 10.

Use the binomial ($n = 17$, $p = .3962$) to calculate

the p-value = $P(X \leq 3) = 0.0495$ (see Minitab calculations below)

Since the p-value is $< .05$, we reject the null H. Conclude that there is sufficient evidence to suggest that Conservative support among students is less than 39.62%.

If the p-value is rounded to 0.05 which is not less than the level of significance, then the decision could be not to reject the null H.

Binomial with $n = 17$ and $p = 0.3962$

x	P(X <= x)
3	0.0495293

Note that Minitab handles this easily if you do not assume a normal distribution for p -hat.

Test of $p = 0.3962$ vs $p < 0.3962$

				99%	
				Upper	Exact
Sample	X	N	Sample p	Bound	P-Value
1	3	17	0.176471	0.479621	0.050

If normal approximation is used, then $z = (3/17 - .3962) / \sqrt{.3962 * .6038 / 17} = -1.85$, but Minitab gives a warning about the small sample size:

Test of $p = 0.3962$ vs $p < 0.3962$

				99%		
				Upper	Z-Value	P-Value
Sample	X	N	Sample p	Bound	Z-Value	P-Value
1	3	17	0.176471	0.391563	-1.85	0.032

* NOTE * The normal approximation may be inaccurate for small samples.

4 marks:

1 for hypotheses

1 for some recognition that the normal approximation is not appropriate

1 for showing a manual calculation of p-value of .05

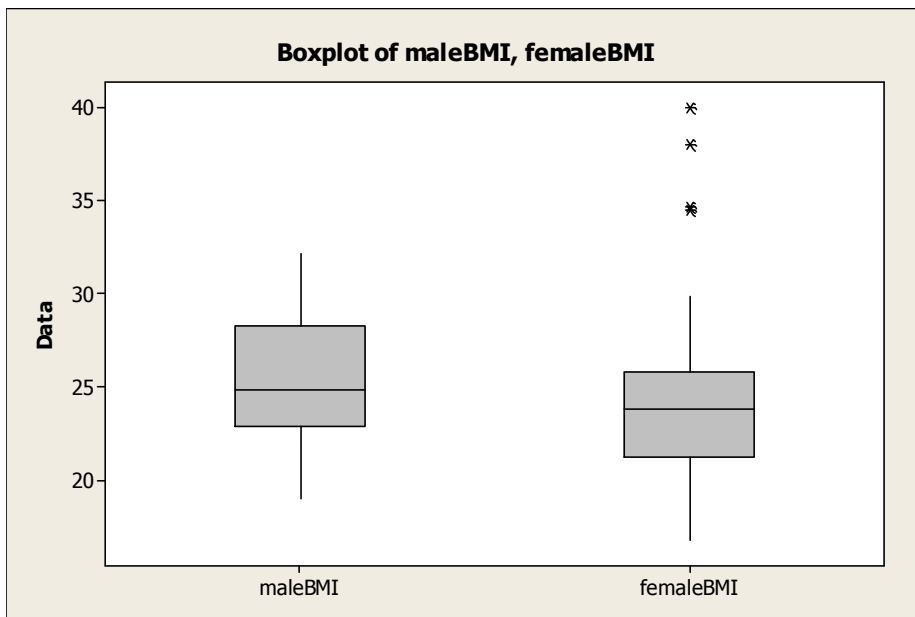
1 for decision and conclusion (if the answer uses the $z = -1.85$, then the decision would be to reject the null H since z is < -1.645 , and the conclusion would be that the Conservative support among students is less than 39.62%.

Thus, if students calculate the z -statistic based on the assumption that the sample proportion is normally distributed and draw the conclusion based on this, then give them 2 marks out of 4.

The sample of students is not representative of the Canadian electorate and is too small to make any inferences about the national level of support for the Conservatives.

[1 mark] for answering the final question.

Problem 3. [10 marks]



Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
maleBMI	39	0	25.749	0.558	3.484	19.000	22.900	24.900	28.300
femaleBMI	34	0	24.706	0.929	5.419	16.800	21.250	23.800	25.800

- (a) Since the samples are large, the CLT assures us that the sample means are normally distributed, provided the population distributions are not extremely skewed. This is reasonable even though there is one extreme outlier (the largest observation of 40 is just beyond 3 IQR beyond the 3rd quartile or 39.45).

However, we will accept the comment that the female BMI sample may be coming from an extremely skewed distribution, thus invalidating the 2-sample t-test.

1 mark: .5 for reasonable comment on the graphs, .5 for reasonable comment on the normality of the sample means and the suitability of the t-test.

As both sample sizes are large, there is no need to assume equal population variances. However, t only changes to .99 if we do pool the sample variances.

Two-Sample T-Test and CI: maleBMI, femaleBMI

Two-sample T for maleBMI vs femaleBMI

	N	Mean	StDev	SE Mean
maleBMI	39	25.75	3.48	0.56
femaleBMI	34	24.71	5.42	0.93

Difference = mu (maleBMI) - mu (femaleBMI)

Estimate for difference: 1.04284

95% CI for difference: (-1.13034, 3.21601)

T-Test of difference = 0 (vs not =): T-Value = 0.96 P-Value = 0.340 DF = 54

Ho: mu(male) - mu(female) = 0

Ha: mu(male) - mu(female) ≠ 0

$$t = (25.75 - 24.71) / \sqrt{3.48^2/39 + 5.42^2/34} = 0.96$$

Since the t-statistic has d.f. exceeding 30, we can find the critical value based on the approximating z distribution.

Do not reject Ho since 0.94 is not > 1.96 at the .05 level of significance.

We cannot conclude there is a difference between the male and female average BMIs.

4 marks:

1 for hypotheses

1 for t-statistic (must show manual calculation)

1 for decision based on critical value or p-value

1 for conclusion

(c) p-value is .34 (not assuming equal variances) or .325 (assuming equal variances)

No manual calculation is necessary.

1 mark for p-value

(d) 2-sided CI is:

$$1.04 \pm 1.96 * \sqrt{3.48^2/39 + 5.42^2/34} = 1.04 \pm 1.96 * 1.084 = 1.04 \pm 2.12$$

Using t based on 54 d.f. would give critical value of 2.00, in which case the

CI is $1.04 \pm 2 * 1.084 = 1.04 \pm 2.168 = (-1.13, 3.21)$

If the pooled standard deviation is calculated at 4.4883,

then the resulting CI is $(-1.05700, 3.14268)$

2 marks for showing manual calculation of CI (do not worry about accuracy)

(e) p-value not $< .05$ and CI covering 0 both lead to non-rejection of null H.

1 mark for noting the same decision and conclusion as before

(f) We could use the Mann-Whitney test a.k.a. Wilcoxon Rank Sum test.

1 mark for identification of the alternative non-parametric test