ADM2304
Winter 2014
Assignment 1 Solutions and Marking Guide [total 40 marks]

1. [14 marks]

(a) [5 marks]

Test and CI for One Proportion

```
Test of p = 0.2962 vs p < 0.2962
                                    99%
                                  Upper
Sample    X     N  Sample p     Bound  Z-Value  P-Value
1       262  1000  0.262000  0.294348    -2.37    0.009

Test of p = 0.3566 vs p < 0.3566
                                    95%
                                  Upper
Sample    X     N  Sample p     Bound  Z-Value  P-Value
1       262  1000  0.262000  0.284872    -6.25    0.000
```

-Ho: p = .2962 Ha: p < .2962
also accept Ho: p = .3566 (10% lower multiplicatively)
-Reject Ho if Z < - 2.33 (.01 level of significance) (-2.32 is acceptable)
-z= (.262 - .2962) / sqrt (.2962 * .7038/ 100) = -.0342 / .0144 = -2.37 *(need manual calculation)*
-Since z = -2.37 < -2.33, we reject Ho at .01 level
-Conclude drop in Conservative vote of more than 10%.

1 mark for each of the above points

(b) [3 marks]

$n = z^{2*} p*(1-p)/M^2 = 2.576^2* .262 * (1 - .262) / .0001 = 12831$

n = 12771 for z = 2.57 or n = 12871 for z = 2.58

Correct answer anywhere between 12771 and 12871

If p = .5 used, then n = 16590
If p = .2962 used, then n = 13834

1 for use of p-hat = .262 to estimate p, 0 for any other value
1 for z = between 2.57 and 2.58
1 for final calculation, as long as it is correctly calculated using the proper formula, even if p or z is wrong.

(c) [6 marks]

Result assuming normal approximation for z:

Test and CI for One Proportion

```
Test of p = 0.3962 vs p < 0.3962

                             99%
                           Upper
Sample  X   N  Sample p     Bound  Z-Value  P-Value
1       2  17  0.117647  0.299434    -2.35    0.009
```

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

Minitab uses the binomial calculation if normal approx. not used.

Test and CI for One Proportion

```
Test of p = 0.3962 vs p < 0.3962

                             99%
                           Upper    Exact
Sample  X   N  Sample p     Bound  P-Value
1       2  17  0.117647  0.409923    0.013
```

p-value = Prob($X \leq 2$) = P(0) + P(1) + P(2) = 0.013, based on Binomial (n=17, p = 0.3962)

Probability Density Function

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

x  P( X = x )
0   0.0001884
1   0.0021021
2   0.0110348
```

Cumulative Distribution Function

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

x  P( X <= x )
2    0.0133253
```

-Ho: p=.3962, Ha: 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=2 is not greater or equal to 10.
-Use the binomial (n=17, p=.3962) to calculate the p-value = P(X<= 2) = 0.013 (see Minitab calculations above) *Worth 2 marks*
-Since the p-value is not < .01, we do not reject the null H. Conclude that there is not sufficient evidence to suggest that Conservative support among students is less than 39.62%.

5 marks:
1 for hypotheses
1 for some recognition that the normal approximation is not appropriate
1 for correct p-value of 0.013
1 for explaining or showing how p-value would be calculated
1 for decision and conclusion (if the answer uses z = -2.35, then the decision would be to reject the null H since z is < - 2.33, and the conclusion would be that the Conservative support among students is less than 39.62%. (0.5 for decision, 0.5 for conclusion)

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 3 marks out of 5 (they would lose 2 for not recognizing problem with z and for wrong p-value)

The sample of students is not representative of the Canadian electorate and is too small to have enough statistical power.

1 mark for any of the above two points..

2. [6 marks]

One-Sample T: BMIfemale

```
Test of mu = 26 vs < 26

                                                95%
                                              Upper
Variable    N     Mean   StDev  SE Mean     Bound      T      P
BMIfemale  40  24.8125  5.4095   0.8553   26.2536  -1.39  0.086
```

-Ho: mu = 26 ; Ha: mu < 26

-T = (24.8 – 26) / .8553 = -1.39 (*manual calculation not required)*

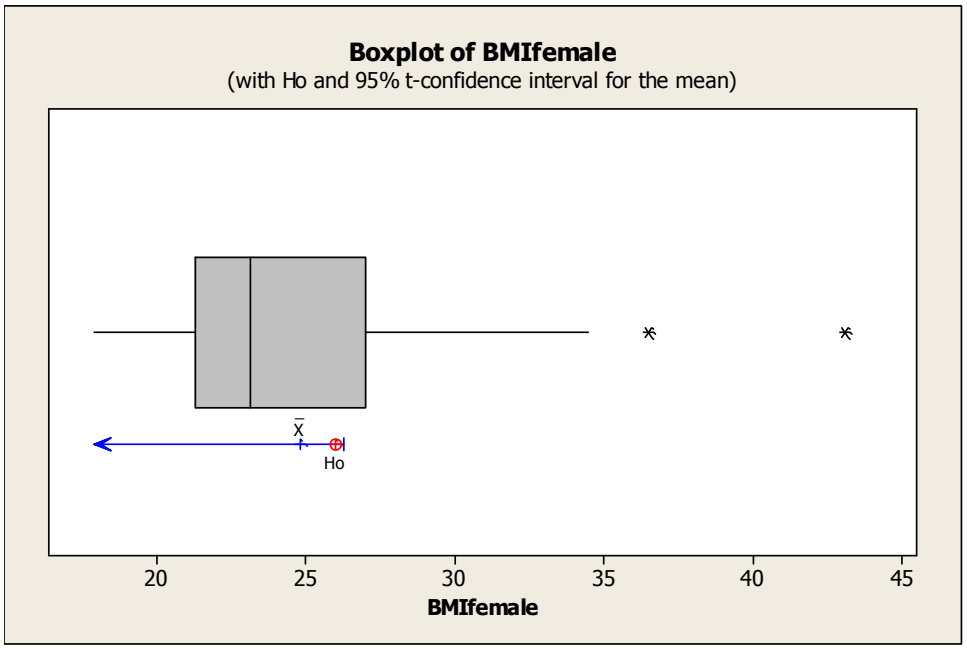
-At .05 level, we reject Ho if t < - 1.645

-Since t is not < -1.645, we do not reject Ho, and Conclude there is insufficient evidence to show average female BMI is less than 26.

1 mark for each of the above four points.

-With a sample size of 40, we need to assume BMI are not extremely skewed. No need to assume a normal distributon. The boxplot shows skewness but not extremely so.

2 marks: 1 for assumption and 1 for comment on this using graphical evidence



3. [9 marks]

(a) [4 marks]

Test and CI for Two Proportions: OWmale, OWfemale

```
Event = 1


Variable   X   N  Sample p
OWmale    25  50  0.500000
OWfemale  11  40  0.275000

Difference = p (OWmale) - p (OWfemale)
Estimate for difference:  0.225
95% CI for difference:  (0.0291568, 0.420843)
Test for difference = 0 (vs not = 0):  Z = 2.17  P-Value = 0.030
```

-Ho: p(OWmale) – p(OWfemale) = 0; Ha: not equal
-calculate pooled proportion of 36/90 = 0.40
-z = .225 / sqrt (.4 * .6 (1/50 + 1/40) = .225 / .104 = 2.16 (*manual calculation must be shown)*
-reject Ho if |Z| > 1.96 (based on 0.05 level) and conclude proportions are different

1 mark for each of the above four points (deduct .5 if no conclusion)

(b) p-value is P(Z > 2.17) + P(Z < -2.17) = 2 * P(Z < -2.17) = 2 * .015 = 0.030

1 mark (must show the probabilities that have to be calculated)

(c) (0.50 – 0.275) +/- 1.96 * sqrt(.5 * .5 / 50 + .275 * .725 / 40)

= (.225) +/- 1.96 * 0.09992 = 0.225 +/- 0.196 = (0.29, 0.421)

2 marks: 0.5 for z value, 1 for standard error formula with proper numbers, 0.5 for final interval. (manual calculation required)

(d) Since the p-value < 0.05 and the 95% CI does not cover zero, we reject the null hypothesis .

2 marks: 1 for each point

4. [11 marks]

(a)

```
Variable    N  N*   Mean  SE Mean   StDev  Minimum    Q1  Median     Q3
Liab07    479   0  39.33     7.57  165.73    0.122  2.18    6.17  17.67
```

One mark for the population mean of 39.33 (units were billion $).

(b) *3 marks for displaying the twenty CIs*

(c) Example of one CI is:

```
Variable   N     Mean    StDev  SE Mean        95% CI
Sample1   50  17.7915  30.8704   4.3657  (9.0183, 26.5648)
```

which does not cover the population mean.

The manual calculation is 17.79 +/- 1.96 * 4.3657
= 17.79 +/- 8.73 = (9.06, 26.52) using z = 2
(could use instead t = 2.01 based on 49 df)

Most boxplots will show skewness, perhaps extremely so, meaning the sampling distribution may not be normal.

1 mark for any graph
1 mark for the assumption that since the sample of 50 is a large sample, the population distribution must not be extremely skewed.
1 mark for any reasonable comment on whether the assumption is reasonable or not (it is probably not reasonable, but accept any reasonable comment since each sample may show more or less skewness.
1 mark for showing manual computation of CI

(d) *1 mark for the count (no need to check the answer)*

(e) The sample mean may not be normally distributed
The sampling fractions exceed 0.05 and no finite population correction was applied by Minitab.
The expected value of 19 is an average over many sets of 20 CI. For any group of 20, random sampling does not guarantee 19 out of 20.

2 marks—1 for each of the above or any others if it is reasonable.