Dr. Suren Phansalker | **ADM2304 X Solution May 28, 2011** | **Assignment#1 115 Marks**

Qu.#1 (15 Marks)

Here:

$\hat{p} = \frac{X}{n} = \frac{657}{750} = 0.8760$, Assume LS = α = 0.05, CC = 1- α = 0.95

S1: 4 Marks

$H_0 : p = (p_0 = 0.9)$

$H_a : p \neq (p_0 = 0.9)$: Two Tailed Test {Not accurate ➔ Less than or More than}

(You must test 'p' and not '$\hat{p}$')

S2: 7 Marks

N.B.: Here,

n p_0 = 50 (0.125) = 6.25 > 5.0, and

n q_0 = 50 (0.95) = 43.75 > 5.0. Thus "Normal" distribution is, strictly speaking reasonable. {3 Marks}

$$\sigma(\hat{p}) = SD(\Pr oportion) = \sqrt{\frac{p_0 q_0}{n}} = \sqrt{\frac{0.9(0.1)}{750}} = 0.0110$$ {2 Marks}

$$z_{Calc} = \frac{\hat{p} - p_0}{\sigma(\hat{p})} = \frac{0.876 - 0.900}{0.0110} = -2.1909$$ {2 Marks}

S5: 4 Marks

p-Val = P[Z > {|z_{Calc} | = 2.1909] x 2

= P[Z > 2.1909] x 2

= (1.0 – 0.9857) x 2

= 0.0286

Since {p-Val = 0.0286} < {α = 0.05}➔ **Reject H0**

You could also use the equivalent critical value approach shown below.

S3:

LS = α = 0.05 and with the Two Tailed Test,

$z_{Crit} = z_{\alpha/2} = z_{0.025} = 1.96$

S4:

Since {|z_{Calc} | = 2.1909} > {z_{Crit} = 1.96} ➔ **Reject H_0. {Just as before}**

Qu.#2 (38 Marks)

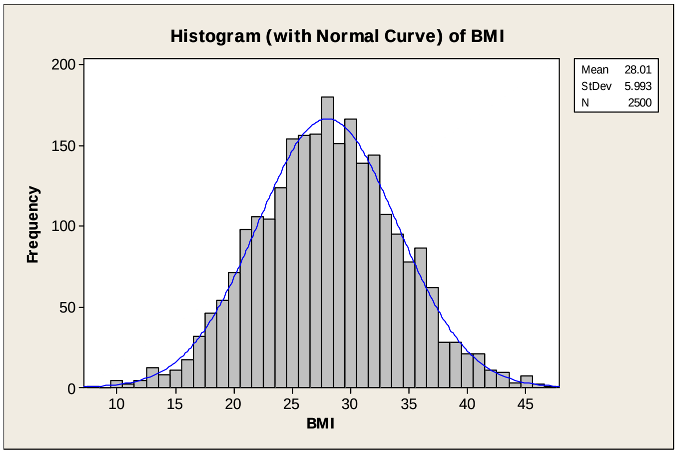
a. -- 6 Marks: 4 for one of the plots, 2 for comments.

Descriptive Statistics: BMI

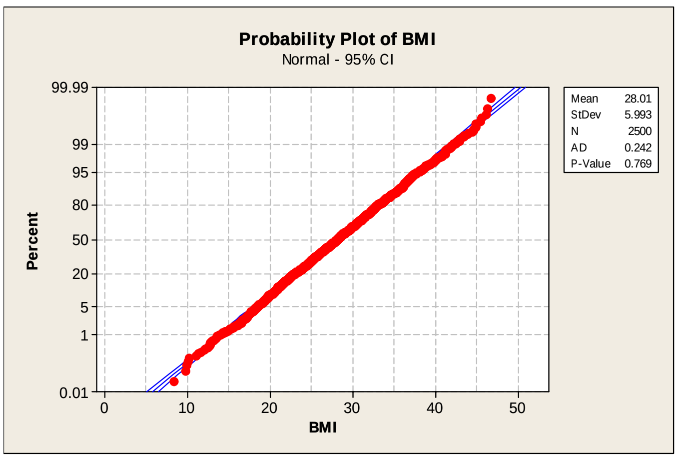
```
Variable      N  N*    Mean  SE Mean  StDev  Minimum      Q1  Median      Q3
BMI        2500   0  28.006    0.120  5.993    8.230  23.960  28.000  32.040

Variable  Maximum
BMI        46.650
```

Histogram (with Normal Curve) of BMI



Or,



Here:
N = 2500
μ = 28.006 units of BMI
σ = 5.993 units of BMI
Here the Histogram as well as the Probability Plot indicates that the population is quite “Normal”. The same can be ascertained by drawing the Boxplot and finding that the population standard deviation, $\sigma \approx \frac{R}{6}$.

b. -- 4 Marks for correct value

Descriptive Statistics: H_Obese

```
Variable      N  N*     Mean  SE Mean    StDev  Minimum       Q1   Median
H_Obese    2500   0  0.12720  0.00667  0.33326  0.00000  0.00000  0.00000

Variable        Q3  Maximum
H_Obese    0.00000  1.00000
```

Since '1' indicates BMI of 35 units or more,
'p' = Population Proportion of people having BMI of 35 or more units is found by simply obtaining the mean of these 2500 '0's and '1's.
There are 318 adults out of 2500 with BMI 35 or more units.

$$p = \frac{318}{2500} = 0.1272$$

c. --8 Marks: 6 for CI, 2 for MiniTab approach.

Here: $n = 50, \bar{X} = 28.752, s = 5.877, CC = 1-\alpha = 0.90, \alpha = 0.10$

$t_{\alpha/2}(n-1) = t_{0.10/2}(49) = t_{0.05}(49) = 1.6766$

CI: $\bar{X} \pm t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} = 28.3318 \pm 1.6766\frac{7.6173}{\sqrt{50}} = 28.3318 \pm 1.8061$

CI: (26.5257, 30.1379)

This CI is verified by using MiniTab as shown below.

One-Sample T: Sample1

```
Variable   N   Mean  StDev  SE Mean       90% CI
Sample1   50  28.33   7.62     1.08  (26.53, 30.14)
```

d. --5 Marks: 4 for CI, 1 for MiniTab approach.

Here:

$\hat{p} = \frac{X}{n} = \frac{8}{50} = 0.1600$, CC = 0.95 = 1- α, and LS = α = 0.05

Thus, $z_{\alpha/2} = z_{0.025} = 1.96$

CI: $\hat{p} \pm z_{\alpha/2}s(\hat{p}) = \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.16 \pm 1.96\sqrt{\frac{0.16(0.84)}{50}}$

CI: 0.16 ± 1.96 (0.0518) = 0.16 ± 0.1016

CI: (0.0584, 0.2616)

CI for One Proportion

```
Sample  X   N  Sample p          95% CI
1       8  50  0.160000  (0.058384, 0.261616)

Using the normal approximation.
```

e. -- 8 Marks: 6 for 20 CIs, 2 for indicating ones which do not contain μ = 28.006.
Obviously, students' 19 other samples will be quite different.

One-Sample T: Sample1, Sample2, Sample3, Sample4, Sample5, Sample6, ...

Variable	N	Mean	StDev	SE Mean	90% CI
Sample1	50	28.33	7.62	1.08	(26.53, 30.14)
Sample2	50	27.564	5.554	0.785	(26.247, 28.880)
Sample3	50	28.766	5.965	0.844	(27.351, 30.180)
Sample4	50	29.426	5.118	0.724	(27.194, 29.791)
Sample5	50	28.783	5.863	0.829	(27.393, 30.173)
Sample6	50	27.66	7.96	1.13	(25.78, 29.55)
Sample7	50	26.473	6.802	0.962	(24.860, 28.086)
Sample8	50	28.522	6.485	0.917	(26.985, 30.060)
Sample9	50	27.180	5.934	0.839	(25.773, 28.587)
Sample10	50	27.311	6.036	0.854	(25.880, 28.742)
Sample11	50	26.222	6.115	0.865	(24.772, 27.672)***
Sample12	50	26.97	7.08	1.00	(25.29, 28.65)
Sample13	50	26.062	6.889	0.974	(24.429, 27.696)
Sample14	50	29.518	5.916	0.837	(28.116, 30.921)***
Sample15	50	28.451	5.066	0.716	(27.250, 29.652)
Sample16	50	27.512	5.884	0.832	(26.117, 28.908)
Sample17	50	27.365	5.539	0.783	(26.052, 28.678)
Sample18	50	28.949	6.794	0.961	(27.338, 30.560)
Sample19	50	27.469	5.916	0.837	(26.066, 28.872)
Sample20	50	28.730	5.314	0.752	(27.470, 29.990)

f.—4 Marks: 2 for the % and 2 for explanation of the expectation.

In the 20 samples, only Sample11 and Sample14 do NOT contain the true value of the population mean μ = 28.006 units of BMI. Obviously, 18 samples contain this true value. {Please Note: The students may obtain a different number.}

Now, $\frac{18}{20} = 0.90$.

90% Confidence Interval suggests that in the "long term", 90 out of 100 (more appropriately 9000 out of 10000, or 90000 out of 100000) confidence intervals would contain the true value of the population mean. Thus, what we obtained was what we had expected although in this small exercise with 20 samples we got some what lucky to get this precise value! {It could very well be different but the explanation about what is expected must be given.}

g.----3 Marks: 1 for recognizing the binomial distribution and 2 for the correct value.

Here X: # of CIs which contain the true Population Mean, μ = 28.006 units
Thus,

$$P[X = x] \sim b(n = 20, p = 0.9, x]$$

$$P[X = (x = 18)] = \binom{n}{x} p^x q^{n-x} = \binom{20}{18}(0.9)^{18}(0.1)^2 = 0.2852$$

Qu.#3 (36 Marks)

a. -- 10 Marks: 2 for each of the 4 steps or their equivalents, and 2 for MiniTab.

S1:

$H_0 : \mu = (\mu_0 = 30.75)$

$H_a : \mu \neq (\mu_0 = 30.75)$:Two Tail Test

S2:

$$s(\bar{X}) = SE(Mean) = \frac{s}{\sqrt{n}} = \frac{7.6173}{\sqrt{50}} = 1.0772$$

$$t_{Calc} = \frac{\bar{X} - \mu_0}{s(\bar{X})} = \frac{28.3318 - 30.75}{1.0772} = -2.2449$$

S3:

LS = α = 0.05

$t_{Crit} = t_{\alpha/2}(n-1) = t_{0.025}(49) = 1.6766$

S4:

Since $\{|t_{Calc}| = 2.2449\} > \{t_{Crit} = 1.6766\}$ ➔ Reject H_0.

The Population Mean, μ, is not 30.75 units of BMI.

p-Val = P{t(49) > |tCalc| = 2.2449} ≈ 0.03

The above results are confirmed by the MiniTab output below.

One-Sample T: Sample1

```
Test of mu = 30.75 vs not = 30.75


Variable   N   Mean  StDev  SE Mean       95% CI          T      P
Sample1   50  28.33   7.62     1.08  (26.17, 30.50)  -2.24  0.029
```

b. – 4 Marks, 1 for 90% CI, 2 for 95% CI and 1 for the explanation of consistency.

One-Sample T: Sample1 : 90% CI {If 90% CI is checked}

```
Variable   N   Mean  StDev  SE Mean       90% CI
Sample1   50  28.33   7.62     1.08  (26.53, 30.14)
```

The 90% CI is identical to the one obtained in Qu. 2 ‘c’.

One-Sample T: Sample1: 95% CI {If 95% CI is checked}

```
Variable   N   Mean  StDev  SE Mean       95% CI
Sample1   50  28.33   7.62     1.08  (26.17, 30.50)
```

The 95% Confidence Interval, CI, given in this MiniTab output is:

CI: (26.17, 30.50)

It can be readily observed that $\mu_0 = 31$ is not contained in this CI. This supports the conclusion reached above.

c. -- 6 Marks: 2 Marks each for S1&S5, 2 for MiniTab values.

Wilcoxon Signed Rank Test: Sample1

```
Test of median = 31.00 versus median not = 31.00

             N for   Wilcoxon            Estimated
          N   Test  Statistic       P       Median
Sample1  50     50      381.5   0.014        27.97
```

S1:

$H_0 : \widetilde{Md} = (\widetilde{Md}_0 = 31)$ (or $Md_{Pop} = 31$)

$H_a : \widetilde{Md} \neq (\widetilde{Md}_0 = 31)$ (or $Md_{Pop} \neq 31$) :Two Tail Test

S5:
LS = α = 0.05
Since {p-Val = 0.014} < { α = 0.05}➔ Reject H_0.
The Population Median is not 31 units of BMI.

d. -- 10 Marks: 2 each for S2, S4 and S5 or their equivalents, and 1 each for S1, S3, MiniTab output, and for observing the fact that Normal approximation is appropriate since $\{np_0 = 6.25\} > 5$ and so is $\{nq_0 = 43.75\}$.

Here:

$\hat{p} = \frac{X}{n} = \frac{8}{50} = 0.1600$, LS = α = 0.05, CC = 1- α = 0.95

S1:

$H_0 : p = (p_0 = 0.08)$

$H_a : p > (p_0 = 0.08)$: Right Tail Test

S2:

$$\sigma(\hat{p}) = SD(\Pr oportion) = \sqrt{\frac{p_0 q_0}{n}} = \sqrt{\frac{0.08(0.92)}{50}} = 0.0384$$

$$z_{Calc} = \frac{\hat{p} - p_0}{\sigma(\hat{p})} = \frac{0.16 - 0.08}{0.0384} = 2.0851$$

N.B.: Here,
n p_0 = 50 (0.125) = 6.25 > 5.0, and
n q_0 = 50 (0.95) = 43.75 > 5.0 {if not 10!}. Thus "Normal" distribution is, strictly speaking reasonable. We will address the issue of 'when distribution is not normal' in Qu.#5.

S3:
LS = α = 0.05 and with the Right Tail Test,

$z_{Crit} = z_{\alpha} = z_{0.05} = 1.645$

S4:
Since $\{z_{Calc} = 2.0851\} < \{z_{Crit} = 1.645\}$ ➔ **Reject H_0.** Or,

S5:
p-Val = P[Z > $\{z_{Calc} = 2.0851]$ x 1
= (1.0 - 0.9815) x 1
= 0.0185

Since {p-Val = 0.0185} < {α = 0.05}➔ **Reject H_0.**

The Population Proportion, p, is likely to be more than 0.08.

Test and CI for One Proportion

```
Test of p = 0.08 vs p > 0.08

                            95% Lower
Sample  X   N  Sample p      Bound  Z-Value  P-Value
1       8  50  0.160000   0.074721     2.09    0.019

Using the normal approximation.
```

e.--- 6 Marks: 2 for the correct formula with z_α (not $z_{\alpha/2}$), 2 for the value and 2 for the explanation.

The appropriate asymmetrical CI for this Right Tail test is the Lower Bound and it is given by:

Lower Bound: $\hat{p} - z_\alpha s(\hat{p}) = \hat{p} - z_\alpha \sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.16 - 1.645\sqrt{\frac{0.16(0.84)}{50}}$

Lower Bound: 0.16 – 0.0853 = 0.0747

The Lower Bound **may not be consistent** with the results of the test because the condition for normal approximation is not satisfied fully.

Qu.#4 (16 Marks)

a. -- 8 Marks: 2 for formula, 2 for 'Z' value, 3 for calculations, and 1 for rounding up.

Here CC = 1 – α = 0.98, α = 0.02, $z_{\alpha/2} = z_{0.01} = 2.326$
σ = 6.00 units of BMI, and margin of error, ME = E = ± 0.75 units of BMI.
Then

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{E^2} = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$$

$$n = \left(\frac{2.326*(6.0)}{0.75}\right)^2 = (18.6080)^2 = 346.2577$$

After rounding up,
n = 347.

b. -- 8 Marks: 2 for formula, 2 for 'Z' value, 3 for calculations, and 1 for rounding up.

Assume a default value of CC = 1 – α = 0.98, α = 0.02, $z_{\alpha/2} = z_{0.01} = 2.326$

$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}}$, and margin of error, ME = E = ± .020

Then

$$n = \frac{z_{\alpha/2}^2 pq}{E^2}$$

But with unknown 'p' and 'q', (p)(q) = 0.5 x 0.05 = 0.25 is the safe maximum value we can use.

$$n = \left(\frac{2.326^2 * (0.5) * (0.5)}{0.020^2}\right) = \left(\frac{2.326 * 0.5}{0.020}\right)^2 = (58.15)^2 = 3381.42$$

n = 3382 {by rounding up}

Qu.#5 (10 Marks)

Here n = 50, $p_0 = 0.01$

n p_0 = 50 (0.01) = 0.5 < 5

n q_0 = 50 (0.99) = 49.5 > 5 {2 Marks for check & comment}

Since both the conditions are not satisfied simultaneously, "Normal" distribution cannot be used. Instead, we must use binomial distribution.

$\hat{p} = \frac{x}{n}$ ➔ $x = 2$ and $X \sim b(n = 50, p = 0.01, x)$

$$
\begin{aligned}
P[X \geq 2] &= 1 - P[X \leq 1] \\
&= 1 - \{P[X = 0] + P[X = 1]\} \\
&= 1 - \left\{\sum_{x=0}^{x=1} \binom{50}{x} (0.01)^x (0.99)^{(50-x)}\right\} \\
&= 1 - \{0.6050 + 0.3056\} \\
&= 1 - 0.9106 \\
&= 0.08944
\end{aligned}
$$

{3 Marks for Calculations and knowing it is the p-Val}

S1:

$H_0 : p = (p_0 = 0.01)$

$H_a : p > (p_0 = 0.01)$: Right Tail Test

S5:

p-Val = P[X ≥ 2]

= 0.08944

Since {p-Val = 0.08944} > {α =.05}➔ Do not Reject H_0.

There is insufficient evidence to reject the null hypothesis; you cannot claim that the population proportion of defective chips is more than 1%.

{4 Marks for S1 & S5}

Here is the MiniTab confirmation without using the normal approximation.

{1 Mark for additional MiniTab output without Normal approximation.}

Test and CI for One Proportion

```
Test of p = 0.01 vs p > 0.01

                              95% Lower    Exact
Sample   X    N  Sample p         Bound  P-Value
1        2   50  0.040000      0.007154    0.089
```

{3 Marks only with this MiniTab output and no manual calculations or S1 & S5}

Aside: Normal approximation is not called for and would have given the following wrong results! The {p-Val = 0.017} < {α = 0.05} making us wrongly reject the H_0.

Test and CI for One Proportion

```
Test of p = 0.01 vs p > 0.01

                              95% Lower
Sample   X    N  Sample p         Bound  Z-Value  P-Value
1        2   50  0.040000      0.000000     2.13    0.017

Using the normal approximation.
The normal approximation may be inaccurate for small samples.
```

{2 Marks only with this MiniTab output and no manual calculations or S1 & S5}