ADM 2304 **Winter 2010**

Solutions to Midterm

Question 1 (10 Marks)

One-Sample T

Many students approached this problem incorrectly, thinking it was a test of proportions. It should have been clear from the boxplot and summary statistics that the data on fund performances were quantitative (albeit measured in percentages) and not qualitative. This is a test of the value of the mean performance of Canadian equity mutual funds (where each fund performance is measured as a percentage), whether it exceeded 25(%). The reference value of 25% is also not a relative count or a proportion, but comes from measure of your portfolio performance).

```
Test of mu = 25 vs > 25
                                95% Lower
 N   Mean  StDev  SE Mean       Bound      T      P
19  30.32   5.59     1.28       28.10   4.15  0.000
```

a. S1

[4] H0: $\mu = 25$

Ha: $\mu > 25$

S2:

Reject H0 if $t > t_{Crit} = t_{0.05}(18) = 1.734$

S3: $t_{Calc} = (X_{-Bar} - \mu_0)/SE(X_{-Bar})$

$= (30.32 - 25)/(5.59/\sqrt{19})$

$= 4.1484$

S4: Since $\{t_{Calc} = 4.1484\} > \{t_{Crit} = 1.734\}$ ➔ Reject H0.

There is statistical evidence to suggest that CEM Funds have a return of more than 25%. (1 Mark for each of the 4 steps).

```
N   Mean  StDev  SE Mean       95% CI
19  30.32   5.59     1.28  (27.63, 33.01)
```

b. 95% Confidence Interval: Symmetrical

[3] X-Bar $\pm\, t_{\alpha/2}(n-1)\ s/\sqrt{n}$ (1 Mark)

$= 30.32 \pm 2.1009\ (5.59/\sqrt{19})$ (1 Mark)

$= 30.32 \pm 2.6943$

$= (27.6257, 33.0143)$ (1 Mark)

Or, one-sided Lower-bound interval (for the Right Tail test)

X-Bar $- t_{\alpha}(n-1)\ s/\sqrt{n} = 30.32 - 1.734\ (5.59/\sqrt{19})$ (1 Mark)

$= 30.32 - 2.2237$ (1 Mark)

$= 28.0963$ (1 Mark)

c. We assume that X-Bar is normally distributed. Since the sample is small, the population from which the (small) sample is drawn must be normal. The boxplot is symmetrical and has no outliers and justifies our assumption.

d. Since the population standard deviation 'σ' is unknown, we can use 's' as a reasonable estimate and use the following formula:
$n = Z^2_{\alpha/2}/\{(ME)^2\}(\sigma)^2 \approx Z^2_{\alpha/2}/\{(ME)^2\}(s)^2 = \{Z_{\alpha/2}/\{(ME)(s)\}^2$ [1 Mark]
[2] $= \{1.96\,(5.59)/1\}^2 = 120.04$
n = 121 (by rounding up) [1 Mark]

In this problem, 1% is not 0.01 but just 1! If you think of the margin of error as .01, then the value of s would be 5.59%, or .0559. You must be consistent in your use of units.

Question 2 (6 Marks)

a. The two sets of before and after data of blood pressure readings are paired through the same patient. This is so because the before and after readings are done for the same 15 patients. In other words, patient#1 had the systolic b.p. of 145 mm of Hg before and 144 mm of Hg after the treatment. The same is true for other 14 patients. Thus the readings are paired.
[1 Mark]

b. Of the three boxplots, the most relevant boxplot is that of the "diff" variable. Although the boxplot is very tiny, it is reasonably symmetrical and it has no real or suspected outliers. Thus the generated sample of differences is reasonably normally distributed and by inference, so is the population it is drawn from. Thus a Parametric t-Test on the "Mean of the Difference" is justified. [1 Mark]

c. S1
[4] H0: $\mu_d = 0$
Ha: $\mu_d > 0$

S2:
$t_{Crit} = t_\alpha(n-1)$
$t_{Crit} = t_{0.05}(14) = 1.7613$

S3: $t_{Calc} = (d_{-Bar} - 0)/SE(d_{-Bar})$
$= (1.6667 - 0)/(2.5542/\sqrt{15})$
$= 2.5572$
S4: Since $\{t_{Calc} = 2.5572\} > \{t_{Crit} = 1.7613\}$ ➔ Reject H0.

There is statistical evidence to suggest that the "Mean of the Difference" in the systolic b.p. of the patients was higher before the treatment and the treatment brought it down. In other words, the new drug is effective.

(1 Mark for each of the 4 steps).

Question 3

a)

1 Mark for hypotheses: $H_0: p = .2$
$H_a: p > .2$

1/2 Mark for estimate: $\bar{p} = \frac{35}{125} = 0.28$

1 Mark for z-statistic: $z = \frac{0.08}{\sqrt{\frac{0.2 * 0.8}{125}}} = 2.24$

1/2 Mark for critical value: 1.645

1 Mark for conclusion: Therefore, there is strong evidence, at the 0.05 significance level, that the true proportion of young adults with health problems is greater than 20%.

b)

P-value = 0.0125

c)

One sided $\left(0.28 - 1.645 \times \sqrt{\frac{0.28 \times 0.72}{125}}\right) = (0.22, 1)$

Two sided $\left(0.28 - 1.96 \times \sqrt{\frac{0.28 \times 0.72}{125}}, 0.28 + 1.96 \times \sqrt{\frac{0.28 \times 0.72}{125}}\right) = (0.21, 0.35)$

1 mark off for every error, full marks for either one-sided or two-sided, although the 1-sided interval is more appropriate.

d)

1 mark for saying there are no assumptions required

Question 4

a)

1 Mark for hypotheses: $H_0: \mu_m - \mu_f = 3$

$H_a: \mu_m - \mu_f > 3$

1 Mark for pooled variance: $s_p^2 = \frac{4.338^2 + 4.186^2}{2} = 18.17$

1 Mark for z-statistic:

z = [(28.285 – 22.440) – 3] / [4.26 * sqrt(1/10)] = 28.45/1.347 = 2.11

1 Mark for critical value: 1.684 or 1.69 (accept either) and decision rule

or 0.025 < p-value < 0.01

1 Mark for conclusion that there is strong evidence, at the 0.05 significance level, that the average male BMI exceeds average female BMI by more than 3.

b)

Normality of BMI distribution in each population – 1 mark

c)

Mann-Whitney U – 1 mark
If your answer is the "Wilcoxon test", you must specify the Wilcoxon Rank Sum test (independent samples test) which is another name for the Mann-Whitney test because the Wilcoxon test is usually the Signed Rank test which is a 1-sample or matched pairs test.