Midterm Fall 2013
Solutions

Question 1.

(a)
-Ho: p = 0.05; Ha: p < 0.05
p-hat = 10/400=0.025
-Z = (0.025 – 0.05) / sqrt(0.05*0.95/400) = -2.29
-rejection region is z < -1.645
-Reject Ho since -2.29 is in the rejection region < -1.645 and conclude the auditor's claim is correct.

1 mark for each point above.

```
Test of p = 0.05 vs p < 0.05

                                    95%
                                  Upper
Sample   X    N  Sample p         Bound  Z-Value  P-Value
1       10  400  0.025000      0.037840    -2.29    0.011
```

(b)
The p-value is Prob(Z < -2.29) = 0.011
Since the p-value is < 0.05, we reject the null H.

1 mark for the correct p-value; & 1 mark for the correct decision and reason.

(c)
The appropriate CI is 1-sided with an upper bound and no lower bound:
UB is 0.025 + 1.645 sqrt (0.025 * 0.975 / 200)
= 0.025 + 1.645 * 0.0078 = 0.025 + 0.013 = 0.038
Interval is (0 , 0.038). Since the CI does not cover the value 0.05, we reject the null H at the 0.05 level of significance.

The 2-sided CI is incorrect but would be
0.025 ± 1.96 * 0.0078 = 0.025 ± 0.015 = (0.10 , 0.040)

Again, since the CI does not cover the value 0.05, we reject the null H.

1 mark for correct 1-sided CI, 0.5 for the 2-sided CI;
1 mark for explaining the decision based on the CI.

(d) n = pq (1.96 / 0.01) ^2.
The best value to use for p would be 0.025, whereupon n = 937 (2 marks for this answer)
Other possibilities would be p = 0.05 or p = 0.5, whereupon n = 1825 or n = 9604, respectively. *(only 1 mark for this answer as these values for p are not reasonable)*

Question 2 [11 marks]

[4] a.

S1: H_0: $\mu = (\mu_0 = 5.5)$, H_a: $\mu \neq (\mu_0 = 5.5)$ ==> Two-tailed test, LS = $\alpha = 0.01$

S2: $SE(X_{Bar}) = s(X_{Bar}) = s/sqrt(n) = 3.452/sqrt(21) = 0.7533$

$t_{Calc} = (X_{Bar} - \mu_0)/SE(X_{Bar}) = (3.443 - 5.5)/0.7533 = -2.7307$

S3: $t_{Crit} = t_{\alpha/2}(n - 1) = t_{0.01/2}(20) = 2.8453$ or 2.85 from the "t" Table

S4: Since $\{|t_{Calc}| = 2.7307\} < \{t_{Crit} = 2.85\}$ ==> Do Not Reject H_0

There is insufficient evidence to claim that the mean return is anything other than 5.5%.

[1]b.

The test in part 'a' is a parametric test on One Population Mean μ. The boxplot of the sample observations indicates that even with the small sample size of 21, there are no outliers and it is reasonably symmetric indicating that the sample and the population it is drawn from is reasonably normal. Thus the parametric "t" test on μ is justified.

[2]c.

p-Val $= P\{t(20) > |t_{Calc}| = 2.7307\} \times 2 \approx \{(0.008 + 0.005)/2\} \times 2 \approx 0.013$

{More precise p-Val is 0.01288}

Since $\{\text{p-Val} \approx 0.013\} > \{\alpha = 0.01\}$ ==> Do Not Reject H_0.

[2]d.

The symmetric Confidence Interval with a Confidence Level (CL) of 99% is found as follows:

$X_{Bar} \pm t_{\alpha/2}(n - 1)s(X_{Bar}) = 3.452 \pm 2.85 \times 0.7533 = 3.452 \pm 2.1469$

99% CI: (1.3051, 5.5989)

{ More precise CI is: (1.3000, 5.5860)}

Since $(\mu_0 = 5.5)$ is in this CI, it is very likely to be 5.5%. Thus the CI results are consistent with the Hypothesis Test which does not reject H_0.

[2]e.

With CL = $1 - \alpha = 0.95$, $\alpha = 0.05$ and $Z_{\alpha/2} = 1.96$, ME = 0.5% and sigma = 3.5%

$n = \{ (Z_{\alpha/2} / ME) \times sigma\}^2$

$n = \{(1.96/0.5 \times 3.5)\}^2$

$n = (13.72)^2 = 188.23$ or $n = 189$

Question 3 [11 marks]

[4]a.

S1: H_0: $\mu_1 - \mu_2 = (\Delta_0 = 0)$, $\mathbf{H_a}$: $\mu_1 - \mu_2 \neq (\Delta_0 = 0)$ **==>** Two-tailed Test with LS = α = 0.05

This is the "Unequal Population Variances" test with "df" = 27 as given in the Minitab output.

S2: $SE(X_{1Bar} - X_{2Bar}) = s(X_{1Bar} - X_{2Bar}) = SQRT\{(s_1)^2/n_1 + (s_2)^2/n_2\}$

$SE(X_{1Bar} - X_{2Bar}) = s(X_{1Bar} - X_{2Bar}) = SQRT\{(3.21)^2/22 + (1.88)^2/47\} = 0.7373$

$t_{Calc} = \{(X_{1Bar} - X_{2Bar}) - (\Delta_0 = 0)\}/ SE(X_{1Bar} - X_{2Bar}) = \{(4.14 - 1.68)/0.7373\}$ = (2.46)/0/7373

$t_{Calc} = 3.3365$

S3: $t_{Crit} = t_{\alpha/2}(df) = t_{0.05/2}(27) = t_{0.025} = 2.05$

{More precise value is 2.05183}

S4: Since $\{|t_{Calc}| = 3.3365\} > \{t_{Crit} = 2.05\}$ ==> Reject H_0

There is sufficient evidence to indicate that the two population means are not equal.

[3]b.

S1: H_0: $|\mu_1 - \mu_2| = (\Delta_0 = 1)$, $\mathbf{H_a}$: $|\mu_1 - \mu_2| \neq (\Delta_0 = 1)$ **==>** Two-tailed Test with LS = α = 0.05. *Note that the absolute values allow us not to specify which population mean is bigger or smaller.*

This is the "Unequal Population Variances" test with "df" = 27 as given in the Minitab output.

S2: $SE(X_{1Bar} - X_{2Bar}) = s(X_{1Bar} - X_{2Bar}) = SQRT\{(s_1)^2/n_1 + (s_2)^2/n_2\}$

$SE(X_{1Bar} - X_{2Bar}) = s(X_{1Bar} - X_{2Bar}) = SQRT\{(3.21)^2/22 + (1.88)^2/47\} = 0.7373$

$t_{Calc} = \{|X_{1Bar} - X_{2Bar}| - 1\}/ SE(X_{1Bar} - X_{2Bar}) = \{((4.14 - 1.68) - 1)/0.7373\}$ = (1.46)/0.7373 = 1.98 *or* $\{(1.68 - 4.14) - (-1)\}/0.7373 = -1.98$

Obviously, $t = \{(1.68 - 4.14) - 1\}/0.7373 = 4.69$ *is incorrect. It is safer to calculate the absolute value of the difference first.*

S3: $t_{Crit} = t_{\alpha/2}(df) = t_{0.05/2}(27) = t_{0.025} = 2.05$

{More precise value is 2.05183}

S4: Since $\{|t_{Calc}| = 1.9802\} < \{t_{Crit} = 2.05\}$ ==> Do not reject H_0

There is *insufficient* evidence to indicate that the two population means are not equal.

[2]c.

The symmetric 95% Confidence Interval is found as follows:

$(X_{1Bar} - X_{2Bar}) \pm t_{\alpha/2}(df)\ s(X_{1Bar} - X_{2Bar})$ = (4.14 - 1.68) ± 2.05 x 0.7373 = 2.46 ± 1.5115, or -2.46 ± 1.5115

95% CI: (0.9485, 3.9715) or (-3.9715, -0.9485)

{ More precise CI is: (0.9470, 3.973)}

Since $(\Delta_0 = 0)$ is in not contained in this CI, it is very likely to be other than 'zero' or the two population means are not likely to be the same. This is consistent with the conclusion of HT in part 'a' above.

Similarly, since (Δ_0 = 1) is contained in (0.9485, 3.9715) and -1 is contained in (-3.9715, -0.9485), the difference is very likely not to be other than 'one' or the two population means are not likely to be different from each other by anything other than 'one'. This is consistent with the conclusion of HT in part 'b' above.

[2]d.

The two boxplots given in Appendix B indicate that both of them have no outliers, and both of them can be considered as being reasonably symmetric and therefore both the samples as well as the populations they are drawn from can be considered as being reasonably normal. *It should be noted that the population that yielded the larger sample need only be not extremely skewed since the sample mean will be normally distributed by virtue of the Central Limit Theorem.*

Finally, since the two age groups are quite unlike each other, the assumption that the population variances are unequal is also appropriate. Thus the parametric "t" test on the difference of means of the two populations with unequal variances is justified.

Question 4.

Assume p = 0.40. *There is no hypothesis test in this question. However, if we have Ho: $p = 0.40$ and Ha: $p < 0.40$, then the desired probability is the p-value for the sample result $X = 175$ (p-hat = 0.35) based on $n = 500$, or $X=1$, based on $n=15$.*

(a)

n = 500, X is Binomial (500, 0.40) with mean 200 and stdev 10.95

$\text{Prob}(X \leq 175) = P ((X - \text{mean})/ \text{stdev} \leq (175 - 200) / 10.95445)$

$= \text{Prob}(Z < -2.28) = 0.0113$

-1 mark for finding the right mean and stdev

-1 mark for finding the right probability of 0.0113

If the correction for continuity is applied, then the 175 becomes 175.5 and we have Prob ($Z < -2.24$) = 0.0125. The correction is not required.

(b)

If n = 15, then we cannot assume a normal distribution for the sample proportion.

Going back to first principles, $P(X \leq 1) = P(X = 0) + P(X=1)$

$= 0.60^{15} + 15 * 0.40 * 0.60^{14} = 0.00047 + 0.0047 = 0.00517$

- *1 mark for recognizing the normal approximation for the binomial is not warranted.*
- *1 mark for calculating the correct binomial probabilities.*