**ADM 2304** **Fall 2012**

**Assignment 2 Solutions**

## Question 2 [ 13 ]

(a) [5]

Ho: p1=p2 or p1 - p2 = 0; Ha: p1≠p2 or p1 – p2 ≠ 0
Pooled p-hat = (7+7)/(77+48)=.112
Z = (.146 - .091)/sqrt(.112*.888*[1/77+1/48]) = .055/.058 = 0.95 or -.95
Reject Ho at .10 level if |z| > 1.645
Do not reject Ho and conclude the failure rates are not different.

*- 1 mark for each of the above parts (must show manual calculation of z)*

### Test and CI for Two Proportions

```
Sample  X   N  Sample p
1       7  77  0.090909
2       7  48  0.145833


Difference = p (1) - p (2)
Estimate for difference:  -0.0549242
95% CI for difference:  (-0.173635, 0.0637860)
Test for difference = 0 (vs not = 0):  Z = -0.95  P-Value = 0.344
```

Alternative solution using 2 x 2 table:

Chi-sq = 0.897 = $(.95)^2$
Some may use the Yates-corrected formula: $\chi 2 = \Sigma i \Sigma j ( |Oij - Eij| -.5 )^2 / Eij = .426$

```
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

          C1     C2  Total
    1      7     70     77
        8.62  68.38
       0.306  0.039

    2      7     41     48
        5.38  42.62
       0.491  0.062

Total     14    111    125

Chi-Sq = 0.897, DF = 1, P-Value = 0.344
```

(b) [4]
Ho: p(A) = 1/7, p(B) = p(C) = p(D or F) = 2/7
Ha: one probability is different from that specified

| A | B | C | D or F | Total |
|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Observed | 20 | 34 | 39 | 32 | 125 |
| Expected | 17.85714 | 35.71429 | 35.71429 | 35.71429 | 125 |
| Chi-sq contributions | 0.257143 | 0.082286 | 0.302286 | 0.386286 | 1.028 |

Chi-square statistic is 1.028
Reject Ho if Chi-square > 6.25 (3 d.f.)
Do not reject Ho, cannot conclude the mark distribution is different from the long-term distribution

*-1 for hypotheses*
*-1 for showing manual calculation of chi-square statistic*
*-1 for rejection region*
*-1 for decision and conclusion*

If the null hypothesis is different, perhaps using five instead of four categories, I would mark it as correct:

Ho: p(A) = p(D) = p(F)= 1/7, p(B) = p(C) = p(D or F) 2/7

| | | |
|---|---|---|
| 20 | 17.85714 | 0.257143 |
| 34 | 35.71429 | 0.082286 |
| 39 | 35.71429 | 0.302286 |
| 18 | 17.85714 | 0.001143 |
| 14 | 17.85714 | 0.833143 |
| 125 | 125 | 1.476 |

Here the chi-square statistic is 1.476 and the critical value is 7.78 (alpha 0.10 and 4 d.f.

A solution with some other set of probabilities that add up to 1 should lose one mark for the hypotheses, but if the rest of the solution makes sense, then give 4 marks out of 5.

(c) [4]
Ho: no overall difference between full-time and part-time students
(student mark and student status are independent)
Ha: some difference between full-time and part-time students
(marks and status are associated or related)

Chi-square statistic is 4.26

Reject Ho if Chi-square > 7.78 (0.10 level of significance)

Do not reject Ho, cannot conclude there is a difference between full-time and part-time students in their marks distribution.

*-1 for hypotheses*

*-1 for showing the manual calculation of chi-square statistic (see details in Minitab output)*
*-1 for rejection region*
*-1 for decision and conclusion*

**Chi-Square Test: C5, C6, C7, C8, C9**

Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

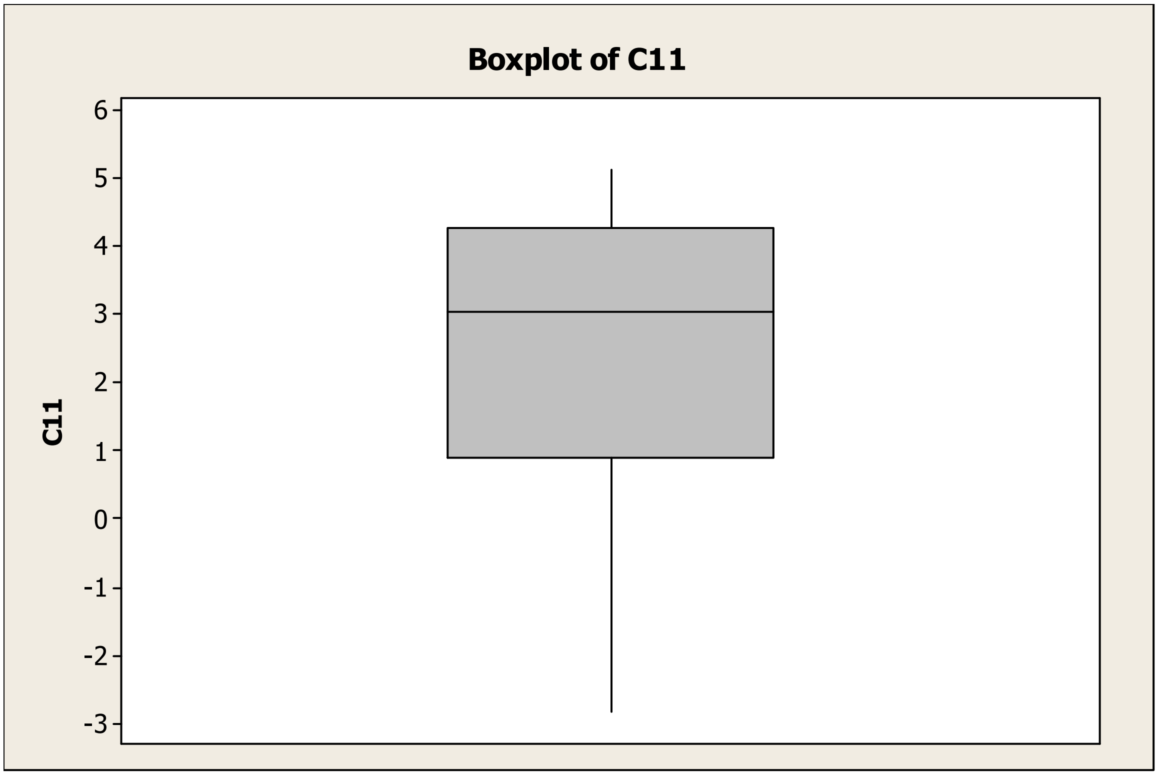
| | C5 | C6 | C7 | C8 | C9 | Total |
|---|---|---|---|---|---|---|
| 1 | 14 | 23 | 25 | 8 | 7 | 77 |
| | 12.32 | 20.94 | 24.02 | 11.09 | 8.62 | |
| | **0.229** | **0.202** | **0.040** | **0.860** | 0.306 | |
| 2 | 6 | 11 | 14 | 10 | 7 | 48 |
| | 7.68 | 13.06 | 14.98 | 6.91 | 5.38 | |
| | **0.367** | **0.324** | **0.064** | **1.380** | 0.491 | |
| Total | 20 | 34 | 39 | 18 | 14 | 125 |

Chi-Sq = 4.261, DF = 4, P-Value = 0.372

**Question 3. [ 10 marks ]**

(a) [ 2 ]
The boxplot is slightly skewed. The nonparametric test is probably safer, but the parametric test is probably acceptable.

*-1 mark for some graph*
*-1 mark for a reasonable comment*




(b) [ 4 ]
Manual computation:

| returns | absval | rank | sign |
|---|---|---|---|
| 4.64 | 4.64 | 9 | 1 |
| 1.37 | 1.37 | 2 | 1 |
| 3.15 | 3.15 | 6 | 1 |
| 3.39 | 3.39 | 7 | 1 |
| -0.47 | 0.47 | 1 | -1 |
| 1.7 | 1.7 | 3 | 1 |
| -2.81 | 2.81 | 4 | -1 |
| 5.14 | 5.14 | 10 | 1 |
| 2.97 | 2.97 | 5 | 1 |
| 4.17 | 4.17 | 8 | 1 |

```
Variable  sign  N  N*     Sum
rank      -1    2   0    5.00
           1    8   0  50.000
```

Ho: median (returns) = 0; Ha: median(returns) > 0

T + = 50, T- = 5

The Wilcoxon statistic is T = 5 as we expect T- to be small under the alternative hypothesis

The rejection region is $T \leq 11$ for a 1-sided test at the .05 level of significance
The rejection region would be larger (a larger critical value for a 1-sided test at the 0.10 level) but the table on Docdepot does not show the 1-sided critical value for alpha = 0.10

Clearly we reject Ho at the .05 level and therefore also at the .10 level
Conclude the median return is positive.

*-1 mark for hypotheses*
*-1 mark for showing how the Wilcoxon statistic is calculated*
*-1 mark for showing rejection region for some alpha level and a 1-sided test*
*-1 mark for decision and conclusion*

(c) [ 2 ]
Same hypotheses as (b)

$$E(T) = n * ( n+1) / 4 = 10(11)/4 = 27.5$$
$$Var(T) = n * ( n+1) * ( 2n+1) / 24 = 10(11)(21)/24= 96.25$$

$Z = [ T – E(T) ] / \sqrt{Var(T)} = (5 – 27.5) sqrt(96.25) = -22.5/9.81 = -2.29$
or $(50 – 27.5) sqrt(96.25) = 22.5/9.81 = -2.29$

We reject Ho if Z positive > 1.645, or if Z negative < -1.645
We reject Ho since the z-statistic is in the rejection region
Conclude the median return is positive.

*-1 mark for calculation of Z-statistic*
*-1 mark for finding the p-value*

Note that the p-value is $P(Z > 2.29)$ or $P(Z< -2.29) = 0.011$

(d) [ 2 ]
Wilcoxon shows T+ = 50.
Using the same hypotheses as (b), we see that the p-value is 0.012 < 0.05
This leads us to reject the Ho at the .10 level and to conclude that the median return is positive.
Note that the p-value using the z-approximation is roughly what Minitab calculated.

- *1 mark for picking out the p-value of -0.012.*
- *1 mark for some comment on the same decision as (b) and (c) or similarity with the p-value in (b).*

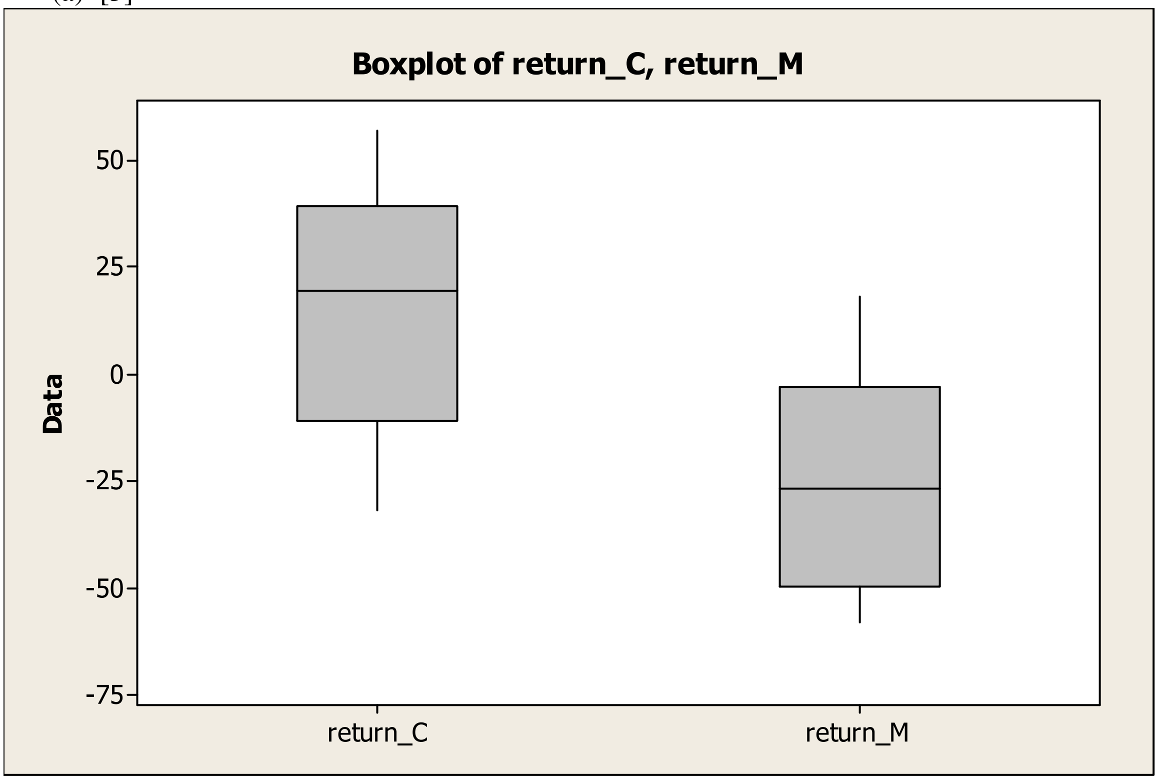
**Wilcoxon Signed Rank Test: returns**

```
Test of median = 0.000000 versus median > 0.000000

                N
              for   Wilcoxon         Estimated
          N  Test  Statistic       P     Median
returns  10    10       50.0  0.012      2.545
```

**Question 4. [ 10 marks ]**

(a) [3]




Both boxplots are relatively symmetric, with no outliers. It is reasonable to assume that both samples come from a normal distribution. The appropriate test is the 2-sample t test.

*-1 mark for some graph showing the two distributions*
*-1 mark for some reasonable comment on the distributions*
*-1 mark for identifying the correct test*

(b) [4]

Ho: mu(C) = mu(M); Ha: mu(C) ≠ mu(M)

Assuming equal pop var, the pooled variance is (5 * 30.7^2 + 7 * 26.2^2)/12 = 793.1275, and the pooled stdev is 28.16
T = 39 / sqrt(793.1275*(1/6+1/8)) = 39 / 15.21 = 2.56
Reject Ho if | t | > 2.18 (.025 point of t with 12 d.f.)

Conclude there is a difference in mean returns between the two types of stocks.

Without assuming equal pop var, we have
t = 39 / sqrt(30.7^2/6 + 26.2^2/8) = 39 / 15.58 = 2.50
Using Minitab to find d.f. = 9, the rejection region is | t | > 2.26
Same decision and conclusion.

*-1 mark for hypotheses*
*-1 mark for showing manual calculation of t-stat*
*-1 mark for the rejection region (no need to calculate the 9 df for the second test)*
*-1 for decision and conclusion*

**Two-Sample T-Test and CI: return_C, return_M**

```
Two-sample T for return_C vs return_M

          N   Mean  StDev  SE Mean
return_C  6   15.5   30.7       13
return_M  8  -23.5   26.2      9.3


Difference = mu (return_C) - mu (return_M)
Estimate for difference:  39.0000
95% CI for difference:  (5.8544, 72.1456)
T-Test of difference = 0 (vs not =): T-Value = 2.56  P-Value = 0.025  DF = 12
Both use Pooled StDev = 28.1684
```

**Two-Sample T-Test and CI: return_C, return_M**

```
Two-sample T for return_C vs return_M

          N   Mean  StDev  SE Mean
return_C  6   15.5   30.7       13
return_M  8  -23.5   26.2      9.3


Difference = mu (return_C) - mu (return_M)
Estimate for difference:  39.0000
95% CI for difference:  (3.7429, 74.2571)
T-Test of difference = 0 (vs not =): T-Value = 2.50  P-Value = 0.034  DF = 9
```

(c) [3 marks]
Ho: median return (C) = median return (M) ; Ha: median return (C) ≠ median return (M)

p-value is 0.06 which is not small enough to reject the null hypothesis

Cannot conclude there is a median difference in the returns

*-1 mark for hypotheses*
*-1 mark for p-value*
*-1 mark for decision and conclusion*

## Mann-Whitney Test and CI: return_C, return_M

```
          N  Median
return_C  6   19.50
return_M  8  -27.00


Point estimate for ETA1-ETA2 is 40.50
95.5 Percent CI for ETA1-ETA2 is (-1.00,77.99)
W = 60.0
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.0612
```