HW1

1. Minimum observation days by vital status

|  |  |
| --- | --- |
|  | Minimum time (days) |
| Did not die | 1827 |
| Died | 68 |

Dichotomizing time to death at 5 years is valid because among those who did not die, the minimum follow-up time of 1827 days is over 5 years (365.25\*5=1826.25 accounting for leap years). Thus, we know the vital status of everyone at 5 years. While there might be censoring, it happens after 5 years of study time.

2. The tale below presents descriptive statistics of the study participants based on time of death (within or after 5 years). LDL levels were higher in the group with death after 5 years. Those who died within 5 years had higher numbers of all other characteristics other than weight.

Characteristics of participants by survival time (Mean (SD) or %)

|  |  |  |
| --- | --- | --- |
|  | Death within 5 years | Death after 5 years |
| Low density lipoprotein (mg/dL) | 119 (36.2) | 127 (32.9) |
| Age, yrs | 76 (6.2) | 74 (5.2) |
| Male | 64% | 47% |
| Weight, lbs | 159 (32.8) | 160 (30.3) |
| Smoking history (pack years) | 28 (36.0) | 18 (24.7) |
| Congestive heart failure | 14% | 4% |
| Coronary heart disease |  |  |
| Angina | 14% | 8% |
| Myocardial infarction | 24% | 10% |
| Stroke |  |  |
| Transient ischemic attack | 6% | 3% |
| Stroke | 23% | 8% |

3. Two sample t-test with unequal variance

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean LDL (mg/dL) | t-statistic | p-value |
| Death within 5 years | 118.7 | -2.38 | 0.02 |
| Death after 5 years | 127.2 |  |  |

Using a two sample t-test with unequal variance, we found that the difference in mean LDL levels (118.7 mg/dL vs. 127.2 mg/dL) for those with death within and after 5 years was statistically significant (t=-2.38, p=0.02). Thus we reject the null hypothesis that the difference in mean LDL levels between groups defined by vital status at 5 years is 0.

4. Geometric mean of LDL by vital status at 5 years

|  |  |  |
| --- | --- | --- |
|  | Geometric mean | 95% CI |
| Death within 5 years | 112.0 | 104.5, 120.0 |
| Death after 5 years | 122.8 | 120.2, 125.5 |
| Difference | 8.8% | 3.4%, 13.8% |

The geometric mean of LDL level in mg/dL for the death was 112.0 (95% CI: 104.5, 120.0) for the death within 5 years group and 122.8 (95% CI: 120.2, 125.5) for the death after 5 years groups. Hence, the geometric mean of LDL levels in the within 5 years group was 8.8% (95% CI: 3.4%, 13.8%) lower than the geometric mean of LDL in the after 5 years group.

5.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Survival within 5yrs (%) | t-statistic | p-value |
| Low LDL | 0.83 | -0.94 | 0.35 |
| High LDL | 0.86 |  |  |

Using a two-sample t-test with unequal variance, we found a 3% higher 5-year survival for the high LDL group. However, there was not a statistically significant difference between the two groups (t=-0.94, p=0.35).

6. The odds of survival past 5 years among those with high LDL is 1.29 times the odds of survival past 5 years among those with low LDL (95% CI: 0.73, 2.28). Since the p-value is above 0.05 (p=0.38), we did not find an association between high LDL level and survival past 5 years.

7. The odds of all-cause mortality among those with high LDL is 0.74 times the odds of all-cause mortality among those with low LDL (95% CI: 0.42, 1.28). As the p-value is greater than 0.05 (p=0.28), we did not find an association between high LDL level and all-cause mortality.

8. Using Kaplan-Meier survival analysis would be best as it does not lose information from categorization and is also appropriate for censored data (since there is loss to follow up after 5 years of study time.