ISSN 0819-2642
ISBN 0 7340 2622 6

THE UNIVERSITY OF MELBOURNE

DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 964

JUNE 2006

Survival on theTitanic: Illustrating Wald and LM Tests for Proportions and Logits

by

Robert Dixon
&
William Griffiths

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

Survival on the Titanic: Illustrating Wald and LM Tests for Proportions and Logits

Robert Dixon and William Griffiths

Corresponding author: Robert Dixon, Department of Economics, University of Melbourne, Victoria 3010, Australia.

Tel: + 61 3 8344 5352; Fax: + 61 3 8344 6899.

E-mail address: r.dixon@unimelb.edu.au

Authors' titles and affiliation

Robert Dixon is a Reader in Economics and William Griffiths is Professor of Econometrics, both in the Department of Economics at the University of Melbourne.

Acknowledgments

The authors are grateful to Bill Becker, Mike Watts and two anonymous referees for their very helpful comments.

Abstract: Students are very interested in lecture examples and class exercises involving data connected to the maiden voyage and the sinking of the liner Titanic. Information on the passengers and their fate can be used to explore relationships between various tests for differences in survival rates between different groups of passengers. Among the concepts examined are tests for differences of proportions using a normal distribution, a chi-square test for independence, a test for the equality of two logits and a test for the significance of the coefficient of a binary variable in logit model. The relationship between Wald and LM test statistics is also examined. Two related examples are given, one to be used for step by step instructional purposes and one to be given as an exercise to students.

Key words: Contingency table, Difference in proportions, Logit model, Statistical tests

JEL codes: A22, A23, C12, C25

On April 10, 1912, the newly completed White Star liner Titanic departed from Southampton on her maiden voyage to New York. At the time she was the largest and most luxurious ship ever built with state-of-the-art equipment and design features that supposedly rendered her unsinkable. At 11:40 PM on April 14, 1912, she struck an iceberg about 400 miles off Newfoundland, Canada. A little less than three hours later, the Titanic plunged to the bottom of the sea, taking almost 1500 people with her.

Information on the characteristics of survivors and non-survivors provides an interesting data set that is convenient for teaching and illustrating a number of statistical concepts.[1] Most students have seen the most recent Hollywood movie about the Titanic and their attention can be readily engaged once they are told the origin of the data. It holds a certain fascination for them, which we can exploit.[2] Also, again thanks to Hollywood, they can readily grasp the hypotheses we will test.

In this paper we use a Titanic passenger data set to test two interesting hypotheses. For the first hypothesis we test whether the survival rate for adult male passengers who were in first class was different from that of other adult male passengers. This issue (the extent to which the ability to reach the lifeboats and thus survive was related to a passenger's socio-economic status or 'class') formed a sub-text to the story line in the most recent Hollywood version of events.[3] To introduce our second hypothesis, we note that some (7%) of the adult males who were traveling in first class cabins were in fact servants in the employ of other first class passengers. If a test of our first hypothesis suggests that there may have been some 'discrimination' between first class and other males in terms of access to lifeboats, etc, it is of interest to test whether this 'discrimination' also existed between the 'upper' and 'lower' classes who were traveling in first class cabins. Thus, for our second hypothesis we ask whether the survival rate for adult male servants traveling in first class was different from that for other adult male first class passengers. While the question of differences in survival rates between first class passengers and others is a long-standing and well-known issue in relation to events on the Titanic, our second hypothesis raises matters not previously considered in the literature surrounding the sinking. Students then can get the feel that they are doing something ‘new’.

The same methodology is used to test both hypotheses, but it is useful to have two related examples; one can be used for classroom presentation and discussion and the other for homework. Also, as we will see, one example involves rejection of a null hypothesis, while, in the other example, the null hypothesis is not rejected. (In our view it is important that students obtain experience with both types of outcomes.) Although the hypotheses are relatively simple, involving only binary variables, they can be used to illustrate a large range of statistical concepts. We describe four ways of testing each hypothesis:

1. A test of the equality of two proportions using a normal distribution.
2. A chi-square test for independence.
3. A test for the equality of two logits.
4. A test for the significance of the coefficient of a binary variable in a logit model.

Describing all four of these tests and the relationships between them improves student understanding of several important concepts and unifies what might seem to be four separate and distinct ways of looking at the same problem. We show how testing the equality of two proportions can be performed as either a Wald test or a Lagrange multiplier (score) test. The chi-square test for independence is shown to be identical to the Lagrange multiplier version of the test. For testing the equality of two logits, we show

how to compute the approximate standard error of a nonlinear function of an estimator. This test can also be performed as a Wald test or a Lagrange multiplier test. We show that a convenient way of computing the test value for the Wald version of the test is to test the significance of the coefficient of a binary variable in a logit model. The fact that different tests for the same problem can lead to different values of test statistics, and potentially different rejection-acceptance decisions, bothers some students. Showing how the same hypothesis can be tested in a number of different ways enhances students' understanding of statistics and helps them grasp the implications of accepting or rejecting a hypothesis.

In the next section the data for testing the equality of the survival rate for adult male passengers traveling in first class with that for other adult male passengers are presented. Subsequent sections are used to describe each of the tests and related material. The results for the hypothesis on survival rates for servants versus other adult male passengers traveling in first class are presented in a final section.

TITANIC SURVIVAL DATA FOR ADULT MALE PASSENGERS

The numbers of adult male passengers on the Titanic, categorized according to class of travel and whether they survived or not, appear in Table 1.[4]

[TABLE 1 NEAR HERE]

The survival rate is defined as the probability that a randomly selected person survived the event. For our purposes it is convenient to assume that the above data represent a random sample from a larger population.[5] Then, if we are considering the survival rate for all adult males, an estimate of this probability is

$$\bar{p} = 131/793 = 0.1652$$

We are interested in comparing the survival rate for adult males traveling first class with that for adult males traveling in the other classes. These rates are conditional probabilities. An estimate of the probability that an adult male survived given that he travelled first class is $\hat{p}_F = 58/176 = 0.3295$; the corresponding probability for adult males traveling in other classes is $\hat{p}_O = 73/617 = 0.1183$. We consider alternative procedures for testing the null hypothesis $H_0 : p_F = p_O$ against the alternative hypothesis $H_1 : p_F \neq p_O$. Or, in other words, we ask whether the difference between $\hat{p}_F = 0.3295$ and $\hat{p}_O = 0.1183$ could be attributable to chance or whether it is indicative of 'discrimination'.

TESTING THE EQUALITY OF TWO PROPORTIONS

Our first test for differences in survival rates is a direct test of the hypothesis $H_0 : p_F = p_O$ against the alternative $H_1 : p_F \neq p_O$. Later tests will be more indirect. For the first test we use the approximate results

$$\hat{p}_F \sim N\left[p_F, \frac{p_F(1-p_F)}{n_F} \right] \qquad \hat{p}_O \sim N\left[p_O, \frac{p_O(1-p_O)}{n_O} \right]$$

where $n_F = 176$ and $n_O = 617$ are the numbers in first and other classes respectively. Assuming the data on first and other classes can be treated as independent samples, the difference between the two sample proportions has the following approximate normal distribution

$$(\hat{p}_F - \hat{p}_O) \sim N\left[(p_F - p_O), \left(\frac{p_F(1-p_F)}{n_F} + \frac{p_O(1-p_O)}{n_O} \right) \right]$$

The corresponding standardized normal random variable, obtained by subtracting the mean and dividing by the standard deviation, is

$$Z = \frac{(\hat{p}_F - \hat{p}_O) - (p_F - p_O)}{\sqrt{\dfrac{p_F(1-p_F)}{n_F} + \dfrac{p_O(1-p_O)}{n_O}}} \sim N[0,1]$$

Two more steps are necessary to convert Z into a test statistic. First, we assume the null hypothesis is true and hence set $(p_F - p_O) = 0$. Second, we need to estimate the standard deviation of $(\hat{p}_F - \hat{p}_O)$ that appears in the denominator of Z. There are two possible estimates of this standard deviation. For the first one p_F and p_O are replaced by their estimates $\hat{p}_F$ and $\hat{p}_O$. In this case we are estimating the standard deviation of $(\hat{p}_F - \hat{p}_O)$ assuming the alternative hypothesis $H_1 : p_F \neq p_O$ is true. The resulting test is called a Wald test. For the second possible estimate of the standard deviation we assume $p_F = p_O$; in this case no difference between the first class and other passengers is assumed and so the survival rate for all adult males ($\overline{p} = 0.1652$) is used to estimate both p_F and p_O. The resulting test is called a Lagrange multiplier or LM test.[6] Denoting the two test statistics by Z_W and Z_{LM}, their formulas and test values for our hypothesis are

$$Z_W = \frac{\hat{p}_F - \hat{p}_O}{\sqrt{\dfrac{\hat{p}_F(1-\hat{p}_F)}{n_F} + \dfrac{\hat{p}_O(1-\hat{p}_O)}{n_O}}} = \frac{0.2112}{0.03774} = 5.597$$

$$Z_{LM} = \frac{\hat{p}_F - \hat{p}_O}{\sqrt{\dfrac{\overline{p}(1-\overline{p})}{n_F} + \dfrac{\overline{p}(1-\overline{p})}{n_O}}} = \frac{0.2112}{0.03173} = 6.656$$

The different treatments of the standard deviation lead to different test statistic values, but both lead to rejection of the null hypothesis.[7] Using a 5% significance level, the two-tailed critical values are ± 1.96. Since $Z_W > 1.96$ and $Z_{LM} > 1.96$, both tests lead us to conclude that the probability of adult male first-class passengers surviving is different from that for other adult male passengers.

The Wald and Lagrange multiplier tests are often expressed as chi-square (χ^2) tests. Because the square of a $N(0,1)$ random variable is a χ^2 random variable with 1 degree of freedom, the χ^2 formulations of the above tests are equivalent tests that lead to identical results. Specifically,

$$\chi_W^2 = Z_W^2 = (5.597)^2 = 31.32 > 3.84$$

$$\chi_{LM}^2 = Z_{LM}^2 = (6.656)^2 = 44.31 > 3.84$$

where $3.84 = (\pm 1.96)^2$ is the two-sided 5% critical value for a χ^2 distribution with 1 degree of freedom. Comparing the square of the test value with the square of the critical value necessarily leads to the same outcome.[8]

TESTING THE INDEPENDENCE OF SURVIVAL RATE AND CLASS

Another way of testing whether the survival rate for first-class adult male passengers is different from that for other adult male passengers is to use a chi-square contingency table test to test for the independence of survival rate and class of travel. It can be shown that this test is identical to the Lagrange multiplier version of the test described in the previous section. To set up the test we begin with the null and alternative hypotheses

H_0 : survival rate and class of travel are independent

H_1 : survival rate and class of travel are not independent

The test statistic is given by

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

where f_o denotes the observed frequencies in each of the four categories (died first class, died other classes, survived first class, survived other classes) in Table 1, and f_e denotes the expected frequencies under the null hypothesis that survival rate and class are independent. The summation is taken over the four categories.

The expected frequency for a given category is given by its row total multiplied by its column total, divided by the overall total. For example, the expected frequency for those adult males who survived and traveled in first class is

$$f_e = \frac{\text{row total} \times \text{column total}}{\text{grand total}} = \frac{131 \times 176}{793} = 29.074$$

The above formula can be motivated in the following way. First, note that the probability of a randomly selected adult male surviving is $\overline{p} = 131/793$ and the probability that a randomly selected male is traveling in first class is $176/793$. If surviving and traveling first class are independent events (H_0 is true), then the probability of a randomly selected adult male being in first class <u>and</u> surviving is equal to the product of these probabilities

$$\frac{131}{793} \times \frac{176}{793} = 0.036664$$

Thus, 0.036664 is the expected proportion of adult males in the survived, first class category. The expected number is

$$f_e = 0.036664 \times 793 = 29.074$$

Note that this calculation is equivalent to that given by the original formula. Similar calculations can be made for the other categories. The expected frequencies appear in parentheses in Table 2.

[TABLE 2 NEAR HERE]

The value of the chi-square statistic is

$$\chi^2 = \frac{(58-29.074)^2}{29.074} + \frac{(73-101.926)^2}{101.926} + \frac{(118-146.926)^2}{146.926} + \frac{(544-515.074)^2}{515.074} = 44.31$$

Note that this value is identical to the χ^2_{LM} value found in the previous section. Although it is not obvious by looking at the two formulas, it can be shown that the test statistics are algebraically equivalent. We also have the same number of degrees of freedom, namely 1. In the general case the degrees of freedom is equal to the number of rows less one multiplied by the number of columns less one. In our case, $(2-1)\times(2-1)=1$. The 5% critical value is 3.84 and, as before, the null hypothesis is rejected very strongly.

TESTING THE EQUALITY OF TWO LOGITS

The logit ℓ of a probability p refers to the transformation

$$\ell = \ln\left(\frac{p}{1-p}\right)$$

where ln(.) denotes natural logarithm and the ratio $p/(1-p)$ is called the odds in favor of an event. Suppose we are interested in the survival of a first class adult male. We saw earlier that an estimate of this probability is $\hat{p}_F = 58/176 = 0.3295$. The corresponding odds is

$$\hat{p}_F / (1-\hat{p}_F) = 0.3295/0.6705 = 0.4915$$

We interpret this value as the probability of surviving relative to the probability of not surviving. An adult male first-class passenger is approximately half as likely to survive as to die. Or, expressed in terms of the inverse, such a passenger is twice as likely to die than to live. The logit of this value is $\hat{\ell}_F = \ln(0.4915) = -0.7102$. Note the bounds of the three different measures. A probability p lies between 0 and 1, the odds $p/(1-p)$ lies between 0 and ∞ and the logit ℓ lies between $-\infty$ and $+\infty$. This characteristic of the logit makes it convenient for further modelling; estimated probabilities derived from it must lie between 0 and 1, as required. An event with a probability of 0.5 has an odds of 1 and a logit of 0. The probabilities of survival and corresponding odds and logits for first class and other adult males are given in Table 3.

[TABLE 3 NEAR HERE]

Suppose, now, that we return to our original null hypothesis $H_0: p_F = p_O$, but instead of expressing it in terms of probabilities, we express it in terms of the corresponding logits, $H_0: \ell_F = \ell_O$ where $\ell_F = \ln[p_F/(1-p_F)]$ and $\ell_O = \ln[p_O/(1-p_O)]$. Clearly, the two hypotheses are equivalent, but, because the logit transformation is a nonlinear one, tests based on estimated logits will be different from those already described.

The first step towards deriving a test is to find an expression for the variance of an estimated logit, such as $\hat{\ell}_F = \ln[\hat{p}_F/(1-\hat{p}_F)]$. For linear functions of random variables of the form $Y = c_1 + c_2 X$, it is taught in first statistics courses that $\text{var}(Y) = c_2^2\,\text{var}(X)$. A generalization of this result that is approximately true for nonlinear functions, say $Y = g(X)$, is

$$\text{var}(Y) = \left(\frac{\partial g}{\partial X}\right)^2 \text{var}(X)$$

Textbooks often refer to this formula as the delta method. Applying it to the logit transformation (for first-class passengers) yields

$$\begin{aligned}\text{var}(\hat{\ell}_F) &= \left(\frac{\partial \ln[p_F/(1-p_F)]}{\partial p_F}\right)^2 \text{var}(\hat{p}_F)\\ &= \left(\frac{1}{p_F(1-p_F)}\right)^2 \times \frac{p_F(1-p_F)}{n_F}\\ &= \frac{1}{n_F\,p_F(1-p_F)}\end{aligned}$$

Carrying out a similar derivation for the other passengers, and assuming independence between $\hat{\ell}_F$ and $\hat{\ell}_O$, the difference between these two estimated logits will have the following approximate normal distribution

$$(\hat{\ell}_F - \hat{\ell}_O) \sim N\left[(\ell_F - \ell_O), \left(\frac{1}{n_F\,p_F(1-p_F)} + \frac{1}{n_O\,p_O(1-p_O)}\right)\right]$$

Proceeding like we did when testing the equality of two proportions, we can set up two test statistics, the Wald statistic Z_W^{ℓ} that uses estimates $\hat{p}_F$ and $\hat{p}_O$ to obtain an estimated standard deviation for the denominator of the test statistic, and the Lagrange multiplier statistic Z_{LM}^{ℓ} that uses $\bar{p}$ to estimate the probabilities in the standard deviation. The formulas for these statistics and the values they yield for our hypothesis are

$$Z_W^{\ell} = \frac{\hat{\ell}_F - \hat{\ell}_O}{\sqrt{\frac{1}{n_F \hat{p}_F (1-\hat{p}_F)} + \frac{1}{n_O \hat{p}_O (1-\hat{p}_O)}}} = \frac{1.2982}{0.2031} = 6.392$$

$$Z_{LM}^{\ell} = \frac{\hat{\ell}_F - \hat{\ell}_O}{\sqrt{\frac{1}{n_F \bar{p}(1-\bar{p})} + \frac{1}{n_O \bar{p}(1-\bar{p})}}} = \frac{1.2982}{0.2301} = 5.642$$

These values are similar, but not identical to those calculated for the null hypothesis $H_0 : p_F = p_O$. Since they are both greater than the 5% critical value of 1.96, they again lead us to conclude that the survival rates for first class and other adult male passengers are different.

TESTING THE SIGNIFICANCE OF THE COEFFICIENT OF A BINARY VARIABLE IN A LOGIT MODEL

In this section we show that the test value $Z_W^{\ell} = 6.392$ can be conveniently calculated as the ratio of an estimated coefficient to its standard error in a logit model estimated from individual record data. To establish this relationship we begin by defining

$$\beta = \ell_F - \ell_O = \ln\left(\frac{p_F}{1-p_F}\right) - \ln\left(\frac{p_O}{1-p_O}\right) = \ln\left(\frac{p_F/(1-p_F)}{p_O/(1-p_O)}\right)$$

The quantity $\frac{p_F/(1-p_F)}{p_O/(1-p_O)}$ is called the <u>odds ratio</u>. It describes the odds of survival for first class adult male passengers relative to those in other classes. If the odds of survival are the same for both classes, it will be equal to 1. In our example the estimated odds ratio is $0.4915/0.1342 = 3.663$; the odds of surviving are approximately 3.7 times greater for adult males traveling first class than for those traveling in other classes. Since the difference between the two logits $\ell_F - \ell_O$ is equal to the log of the odds ratio, we can test whether the estimated odds ratio is significantly different from one by testing whether $\hat{\beta} = \hat{\ell}_F - \hat{\ell}_O$ is significantly different from zero. In the previous section we found that $\hat{\beta} = 1.2982$ (notice the numerator in the calculations for Z_W^{ℓ} and Z_{LM}^{ℓ}). This same estimate can be computed as a coefficient in an estimated logit model.

Let $\ell_i = \ln[p_i/(1-p_i)]$ be the logit for the i-th adult male passenger, with p_i being the probability that the i-th adult male passenger survived. Since we have 793 adult male passengers, we have $i = 1, 2, ..., 793$. Some of these passengers (176 of them) traveled first class, the others (the remaining 617) traveled in other classes. A logit model relating the probability of survival to class of travel can be written as

$$\ell_i = \alpha + \beta X_i$$

where X_i is a dummy (binary) variable that denotes class of travel. It is equal to 1 when the i-th observation is a first-class passenger and 0 when the i-th observation is a traveler from another class. The parameter α is equal to the logit for "other-class" passengers. That is, $\alpha = \ell_O$. To prove that the above model is equivalent to our earlier formulation, note that substituting $\alpha = \ell_O$ and $\beta = \ell_F - \ell_O$ into above equation yields

$$\ell_i = \ell_O + (\ell_F - \ell_O)X_i$$

When $X_i = 1$, we have $\ell_i = \ell_F$ and when $X_i = 0$, we have $\ell_i = \ell_O$. Thus, by estimating the logit model $\ell_i = \alpha + \beta X_i$ we find

1. An estimate of the intercept parameter α is an estimate of the log-odds of survival for the other-class adult males.
2. An estimate of the coefficient β is an estimate of the log of the odds ratio for survival of first-class adult males relative to other-class adult males.
3. An estimate of the sum $\alpha + \beta$ is an estimate of the log-odds of survival for first-class adult males.
4. Testing $H_O : \ell_F = \ell_O$ is equivalent to testing $H_O : \beta = 0$.

To estimate the logit model we use observations (Y_i, X_i), $i = 1, 2, ..., 793$, for each of the adult male passengers. The variable Y_i is a binary variable equal to 1 if the i-th passenger survived and 0 if the i-th passenger did not survive; X_i is 1 for first-class passengers and 0 otherwise. The data set can be easily created using the information given in the contingency table (Table 1). The total number of observations on each variable is 793. The Y variable will consist of 131 ones followed by 662 zeros. The X variable will consist of 58 ones and 73 zeros (corresponding to the 131 values where $Y = 1$) followed by 118 ones and 544 zeros (corresponding to the 662 values where $Y = 0$).

The maximum likelihood estimates of α and β and related information obtained using the software EViews are given in Table 4.[9]

[TABLE 4 NEAR HERE]

Note that the estimates for α and β agree with the estimates that we obtained earlier using estimated proportions. Specifically,

$$\hat{\alpha} = \hat{\ell}_O = -2.008 \qquad\qquad \hat{\beta} = \hat{\ell}_F - \hat{\ell}_O = -0.710 - (-2.008) = 1.298$$

and

$$\hat{\alpha} + \hat{\beta} = -2.008 + 1.298 = -0.710 = \hat{\ell}_F$$

Furthermore, the standard error for $\hat{\beta}$, 0.203, is identical to the estimated standard deviation in the denominator of the Wald statistic Z_W^{ℓ}. Given that the numerator in this statistic is equal to $\hat{\beta} = 1.298$, it follows that the *Z*-value in the above table is equal to Z_W^{ℓ} and the test for significance of $\hat{\beta}$ from the logit model is equivalent to the Wald test used earlier to test whether the difference between the two logits $\hat{\ell}_F - \hat{\ell}_O$ is significantly different from zero.[10]

Finally, we note that all of the tests we have performed lead us to conclude that the survival rate for male passengers traveling in first class was different from the survival rate for males traveling in other classes on the Titanic.

A SECOND EXAMPLE AND A NEW HYPOTHESIS CONCERNING SURVIVAL RATES ON THE TITANIC

Working through our first example students will have seen that the survival rate for males traveling in first class cabins was significantly higher than the survival rate for males who were not traveling first class. Now, as mentioned in the introduction, some (7%) of the adult males who were traveling in first class cabins were in fact servants (they are described in the Titanic’s passenger list as a ‘servant’, ‘butler’, ‘valet’, ‘clerk’ or ‘secretary’) in the employ of other first class passengers. Given this information, an obvious follow-up to the first example is to investigate whether the servants in first class, like the ‘lower-class’ passengers traveling second or third class, had a different (and lower) survival rate than the adult males in first class who were not servants. Indeed, computing survival rates from the raw data it is the case that the survival rate for servants in first class cabins was 2/12 = 0.167 while the survival rate for other males traveling in first class cabins was 56/164 = 0.342. At first sight this is an important difference, consistent with there having been discrimination not only between passengers traveling in different classes of cabin but also between socio-economic groups within the group of passengers traveling in first class cabins. This potential discrimination is the motivation for our second example where we restrict our attention to adult males who were traveling in first class cabins, and ask whether the survival rate for servants traveling in first class is sufficiently different from that for other adult male first class passengers to reject the null that the difference may have arisen simply due to ‘chance’.

There are a number of ways an instructor could proceed with this example. Students could be given the data on the numbers in each of the four categories (servant and non-servant, survived and did not survive) and asked to work through some or all of the tests that we describe, namely, testing the difference between two proportions, the chi-square test for independence, and testing the difference between two logits. Alternatively, they could be given the individual record data and asked to estimate a logit model, interpret the estimates and then test for differences in survival rates. If this strategy is adopted, a thorough understanding of the material delivered in the first example could be assessed by asking students to use the estimated logit coefficients to find the odds ratio, the odds in favor of survival for both categories, and estimates of the probability of survival for servants and non-servants. That is the approach we follow here. For instructors keen to use the data for students to work through some or all of the other tests that we described in previous sections we present the relevant information in an appendix to this paper.

There were 176 adult males traveling in first class cabins. Of these 12 were servants. The individual record data can be arranged in two columns. The first column (Y) is scored 1 if the passenger survived and 0 if he did not. The second column (X) is scored 0 if the passenger was a servant who was in a first class cabin and 1 if the passenger was traveling first class but was not a servant. As mentioned, the total number of observations on each variable is 176. The Y variable will consist of 58 ones followed by 118 zeros. The X variable will consist of 56 ones and 2 zeros (corresponding to the 58 values where $Y=1$) followed by 108 ones and 10 zeros (corresponding to the 118 values where $Y=0$).

Estimating the logit model for this data yields the results reported in Table 5.

[TABLE 5 NEAR HERE]

Students should be able to explain in words what the two regression coefficients separately and together tell them about the odds of survival for the two types of first-class male passengers (those who were employed as servants and those who were not servants). Specifically, students should report that the constant (-1.6094) is an estimate of the log of the odds of survival for the servants and that the slope coefficient (0.9527) is an estimate of the logarithm of the ratio of the odds of survival for the two groups of (male) passengers traveling in first class cabins. They should also report that the sum of the slope and constant (-0.6568) yields an estimate of the log of the odds of survival for the first class male passengers who were not servants.[11] The anti-logarithms of these figures are 0.2000, 2.5926 and 0.5185 respectively. The students should then be asked to explain in words what each of the numbers represents (is an estimate of) in relation to this specific example. Applying the rules set out earlier for the interpretation of logit results they should report that the estimate of the odds of survival for the servants is 0.200, the estimate of the ratio of the odds of survival for non-servants relative to servants for (male) passengers traveling in first class cabins is 2.593 and the estimate of the odds of survival for the first class male passengers who were not servants is 0.519. Students could also be asked to go one step further and recover the underlying probabilities of survival once they have computed the odds. Since the *odds* = $\left[p/(1-p)\right]$, it follows that p = $\left[odds/(1+odds)\right]$. Applying this rule, students should report that their estimate of the probability of survival for the servants is 0.167 and their estimate of the probability of survival for the first class male passengers who were not servants is 0.341.

Once the task involving estimation and/or interpretation has been completed, students should be asked to use the output from the logit model to test the null hypothesis that the odds of survival are, in truth, the same regardless of whether the passenger traveling in first class was a servant or not a servant. They should also be asked to state in words what conclusions they draw from the test result. If we form a 95% confidence interval around the estimate of β (*b*), that is, we compute $b \pm 1.96\left(S_b\right)$, we find that the interval is (-0.559, 2.505), a range which includes zero; thus, we cannot reject the null that the odds are the same for the two groups and that the apparent difference is arising due to chance.[12] In other words, despite the fact that the odds ratio is $0.5185/0.2 = 2.593$ (the odds of surviving for non-servants is 2.6 times greater than that for servants), we are unable to show statistically that this ratio is significantly greater than 1. It is more difficult to establish statistical significance when sample size is small - and the number of servants was relatively small. (We know in this case that the small sample size is creating a problem for us for the following reason: If, instead of asking whether the survival rate for servants traveling in first class is different from that for other adult male first class passengers, we were to ask whether the survival rate for servants traveling in first class is different from that for male passengers traveling in second and third classes (taken together),[13] we would accept the null that there was no difference. So we would be in a position where we would accept that first class passengers had a higher survival rate than second and third class passengers (taken together) and we would accept that the survival rate for servants traveling first class was no different from that for other males traveling first class, while at the same time accepting the null that the survival rate for servants traveling first class was no different from that for males traveling second or third class (taken together). These outcomes illustrate the 'full' import of a small sample size – the results do not lead us to support one hypothesis or another, but imply we can draw no firm conclusions on the basis of statistical tests.)

SUMMARY

In this paper we used a Titanic passenger data set to test two interesting hypotheses: (i) whether the survival rate for adult male passengers who were in first class was different from that of other adult male passengers, and; (ii) whether the survival rate for adult male servants traveling in first class was different from that for other adult male first class passengers. Testing these hypotheses allowed us to examine a number of statistical concepts including tests for differences of proportions using a normal distribution, a chi-square test for independence, a test for the equality of two logits and a test for the significance of the coefficient of a binary variable in logit model. The relationship between Wald and LM test statistics was also examined, and the delta method for finding the standard error of a nonlinear function of an estimator was illustrated.

APPENDIX

CONTINGENCY AND OTHER TABLES DEALING WITH THE TWO TYPES OF FIRST CLASS PASSENGERS

This appendix refers to our second example where we restrict our attention to adult males who were traveling in first class cabins, and ask whether the survival rate for servants traveling in first class is different from that for other adult male first class passengers. In the main text we looked at the estimation of a logit model using the individual record data. In addition, or as an alternative students could be given the data in the form of a contingency table and asked to test the difference between two proportions, perform a chi-square test for independence and test the difference between two logits. We present the relevant information below.

There were 176 adult males traveling in first class cabins. Of these 12 were servants. Table A1 below gives the number of survivors and non-survivors for each group.

[TABLE A1 NEAR HERE]

We are interested in comparing the survival rate for male servants traveling first class with that for other adult males traveling first class. Using the information provided in Table A1 we find that an estimate of the probability that a male servant survived given that he travelled first class is 0.1667; the corresponding probability for all other adult males traveling first class is 0.3415. The Wald and LM tests statistics for testing the equality of these two proportions are $Z_W = 1.536$ and $Z_{LM} = 1.244$. Using a 5% significance level both test statistics lead us to accept the null of no difference in the two proportions.

Another way of testing whether the survival rate for first-class adult male passengers is different from that for other adult male passengers is to use a chi-square contingency table test to test for the independence of survival rate and whether the passenger was a servant or not. Actual and expected frequencies are given in Table A2. The value of the chi-square statistic is $\chi^2 = 1.546$. With 1 degree of freedom the 5% critical value is 3.84 and, as with the Wald and LM tests, the null hypothesis is accepted.

[TABLE A2 NEAR HERE]

Table A3 reports the probabilities, the odds and the corresponding logits.

[TABLE A3 NEAR HERE]

The test for equality can be performed in terms of the corresponding logits, $H_0 : \ell_F = \ell_O$ where $\ell_F = \ln[p_F/(1-p_F)]$ and $\ell_O = \ln[p_O/(1-p_O)]$.

Proceeding like we did when testing the equality of two proportions, we can set up two test statistics, the Wald statistic Z_W^ℓ that uses estimates $\hat{p}_F$ and $\hat{p}_O$ to obtain an estimated standard deviation for the denominator of the test statistic, and the Lagrange multiplier statistic Z_{LM}^ℓ that uses $\overline{p}$ to estimate the probabilities in the standard deviation. The values of these statistics for this hypothesis are $Z_W^\ell = 1.203$ and $Z_{LM}^\ell = 1.497$. Since they are both less than the 5% critical value of 1.96, they also lead us to conclude that the survival rates for servants and others traveling in first class are not different.

REFERENCES

Baltagi, B. H. 2002. *Econometrics*. Third edition. Berlin: Springer.

Becker, W. E. 1998. Engaging students in quantitative analysis with the academic and popular press. In *Teaching economics to undergraduates: Alternatives to chalk and talk,* ed. W. E. Becker and M. Watts. Cheltenham: Edward Elgar.

Bera, A. K. and G. Premaratne. 2001. General hypothesis testing. In *A companion to theoretical econometrics*, ed. B. H. Baltagi. Oxford: Blackwell.

Dawson, R. J. 1995. The `unusual episode' data revisited. *Journal of Statistics Education* [online], 3(3). Available from World Wide Web: (http://www.amstat.org/publications/jse/v3n3/datasets.dawson.html)

Kennedy, P. 2003. *A guide to econometrics*. Cambridge, Mass: MIT Press.

Leet, D. and S. Houser. 2003. Economics goes to Hollywood: Using classic films and documentaries to create an undergraduate economics course. *Journal of Economics Education* 34 (Fall): 326-32.

Mitcham, L. 2001. ‘RMS Titanic: Passenger Numbers: The Statistics of the Disaster’, Research article from the *Encyclopedia Titanica* [online]. Available from World Wide Web: (http://www.encyclopedia-titanica.org/articles/statistics_mitcham.shtml)

Moore, D. S 2004. *The basic practice of statistics*, W H Freeman and Co, New York.

Simonoff, J. S. 1997. The ‘unusual episode’ and a second statistics course. *Journal of Statistics Education* [online], 5(1). Available from World Wide Web: (http://www.amstat.org/publications/jse/v5n1/simonoff.html)

Watts, M. 1998. Using literature and drama in undergraduate economics courses. In *Teaching economics to undergraduates: Alternatives to chalk and talk,* ed. W. E. Becker and M. Watts. Cheltenham: Edward Elgar.

NOTES

1. We are not the first to realise that this event attracts students' attention. Dawson (1995) and Simonoff (1997) use cross-tabulations and logistic regression to relate the Titanic's passengers survival rate to age, gender and economic status. To maintain student interest, they present their results without disclosing the nature of the disaster, and ask students to try to identify the historical event which put these people at risk. Moore (2004, p 559) uses passenger survival rates in one of a number of revision questions at the end of a chapter concerning the use of chi-squared tests.

2. Becker (1998), Watts (1998) and Leet & Houser (2003) all urge economics (and econometrics) educators to use of examples from 'the world around us', pointing out that they are a powerful means by which to motivate students and to encourage discussion and participation.

3. We focus solely on the adult *male* passengers because the women-and-children-first policy meant that virtually all females traveling in first class were able to reach a lifeboat and survive, whether they were a servant or not. We concentrate solely on adults because a comparison of the results from our two hypotheses is more relevant if male children are excluded.

4. These figures differ slightly from those given in the Dawson (1995), Simonoff (1997) and Moore (2004). There are two reasons for the differences. First, the data we use are taken from Mitcham (2001) and are more accurate, reflecting more recent research on the passengers and the (various) passenger lists. Second, because of our interest in the survival rate of the first class passengers' servants

(all of whom were adults), we are only looking at the data for adults whereas Dawson, Simonoff and Moore provide figures which include children as well as adults.

5. All of the statistical techniques we wish to teach are based on the assumption that we have random samples from a larger population. However, it could be said that we have the complete population (everybody aboard the Titanic). We hope that this issue can be overlooked here given our aim is to find a data set that will ‘grab’ the students’ attention and use it as an aid to teaching them techniques they can use elsewhere. A good classroom exercise might be to consider whether there is any way in which the data can be thought to come from a ‘random sample’. (Perhaps the passenger manifest can be seen as ‘incomplete’ and as a sample drawn from a larger population of candidate groups who might have been on the ship that night.)

6. Students should note that the LM test uses an estimate of variance under the null whereas the Wald test uses an estimate of variance under the alternative. Since the null is more ‘restrictive’ than the alternative, we say that the Wald test is based on the unrestricted estimator while the LM test (invented by C. R. Rao, and sometimes called ‘Rao’s (efficient) score test’) is based on the restricted estimator. There is a nice story which might help students remember the difference between the two measures. The great statistician Ronald Fisher once invited Abraham Wald and C R Rao to his house for afternoon tea. “During their conversation, Fisher mentioned the problem of deciding whether his dog, who had been going to an ‘obedience school’ for some time was disciplined enough.” After a moments reflection Wald, who “lost his family in the concentration camps and

was averse to any restrictions, simply suggested leaving the dog free and seeing whether it behaved properly". Rao reflected on Wald's answer and suggested an alternative test. Rao, who "had observed the nuisances of stray dogs in Calcutta streets, did not like the idea of letting the dog roam freely and suggested keeping the dog on a leash at all times and observing how hard it pulls on the leash. If it pulled too much it needed more training." Adapted from Bera and Premaratne (2001, p 58).

7. These versions of the Wald and LM test statistics both have the same $N(0,1)$ asymptotic distribution under the null hypothesis, but they will yield different values unless $\hat{p}_F = \hat{p}_O$. Baltagi (2002, p 30) has an example of the difference between the Wald and LM tests for testing a hypothesis about a sample proportion. Kennedy (2003, Ch 4) has a good discussion of the circumstances under which one or the other test might be appropriate.

8. Of course, the chi-square formulations are more general. They can be used when the null hypothesis is a joint hypothesis involving more than one equality. The correspondence we have described only holds for null hypotheses with a single equality.

9. The values given in Table 4 will be obtained from any software that uses second derivatives of the log-likelihood function to estimate the information matrix.

10. It can also be shown algebraically that the two alternative estimators and their standard errors are identical. The advantage of using the estimated logit model is the ease with which the results can be calculated using most statistical software.

11. Clearly, -1.6094 + 0.9527 = -0.6567, not -0.6568. However, -0.6568 is the value obtained by considering more decimal places. Here, as elsewhere, we have chosen to report the values that are free from rounding error.

12. This outcome is supported by the results of the other tests that are listed in the appendix. Students could be encouraged to explore the relationships between these results.

13. It will be recalled that earlier we found the survival rate for male first class passengers was significantly different from that for males traveling in the other two classes taken together. The data for testing hypotheses comparing the survival rate for male servants traveling in first class with the survival rate for male passengers traveling in second and third classes taken together can be found in Tables 1 and A1.

TABLE 1

Contingency table for adult male passengers on the Titanic

	First class	Other classes	Row totals
Survived	58	73	131
Did not survive	118	544	662
Column totals	176	617	793

TABLE 2

Actual and expected frequencies for adult male passengers on the Titanic

	First class	Other classes	Row totals
Survived	58 (29.074)	73 (101.926)	131
Did not survive	118 (146.926)	544 (515.074)	662
Column totals	176	617	793

TABLE 3

The probabilities of survival and corresponding odds and logits for first class and other adult male passengers

	First class	Other classes
Probability of survival (p)	0.3295	0.1183
Probability of dying $(1-p)$	0.6705	0.8817
Odds $[p/(1-p)]$	0.4915	0.1342
Logit $\ln[p/(1-p)]$	-0.7102	-2.0085

TABLE 4

Maximum likelihood estimates of α and β: First class versus other classes

Coefficient	Estimate	Std. Error	*Z*-value	*p*-value
α	-2.0085	0.125	-16.113	0.000
β	1.2982	0.203	6.392	0.000

TABLE 5

Maximum likelihood estimates of α and β: Servants versus non-servants

Coefficient	Estimate	Std. Error	*Z*-value	*p*-value
α	-1.6094	0.775	-2.077	0.038
β	0.9527	0.792	1.203	0.229

TABLE A1

Contingency table for the two types of first class male passengers

	Non-servants	Servants	Row totals
Survived	56	2	58
Did not survive	108	10	118
Column totals	164	12	176

TABLE A2

Actual and expected frequencies for the two types of first class male passengers

	Non-servants	Servants	Row totals
Survived	56 (54.045)	2 (3.955)	58
Did not survive	108 (109.95)	10 (8.045)	118
Column totals	164	12	176

TABLE A3

The probabilities of survival and corresponding odds and logits for the two types of first class male passengers

	Non-servants	Servants
Probability of survival (p)	0.3415	0.1667
Probability of dying $(1-p)$	0.6585	0.8333
Odds $[p/(1-p)]$	0.5185	0.2000
Logit $\ln[p/(1-p)]$	-0.6568	-1.6094