Hutcheson, G. D. and Shaefer, L. (2012). Test selection in the 21st century. Journal of Modelling in Management. 7, 3: 375-387.
http://www.emeraldinsight.com/loi/jm2
http://dx.doi.org/10.1108/jm2.2012.29707caa.002

Journal of Modelling in Management

Graeme D. Hutcheson
Les Schaefer

Test selection in the 21st century.

This tutorial demonstrates how a regression-based analytical system can replace the more traditional cook-book approach that still commonly forms the basis of many introductory statistical courses in the social sciences. Analyses based on generalized linear models (GLMs) offer significant advantages in terms of their flexibility and power and also offer a much simpler system for teaching; one that is based on a solid theoretical base. The huge selection of hypothesis tests that are available may be represented using, essentially, a single model, which can be applied to different types of data and study design. Comparison analyses are provided in this tutorial to illustrate the superiority of a regression-based approach. It is proposed that regression models should replace the out-dated hypothesis-testing approach that for some reason still seems to dominate the social sciences.

It is very common in social science research to compare groups in order to ascertain whether statistically significant differences exist. For example, ‘do male and female employees have different perceptions about promotion prospects?’, ‘is educational level related to attitude?’, ‘has a certain intervention strategy affected production output?’ and ‘is the socio-economic status of parents related to their children's physical health?’. Selecting an appropriate statistic to test such hypotheses is a basic analytical skill, but is one that frequently causes problems, particularly for students who rely on ‘traditional’ tests such as the t-test, ANOVA, Mann-Whitney, Kruskal-Wallis, Friedman and the chi-square tests for contingency tables. There are a bewildering number of tests available and selection is often the result of using simple tables or flow-charts on the basis of decisions about the number of variables involved, the number of groups compared, the design of the study and the level of measurement of the dependent variable (see, for example, Motulsky 2010 and Motulsky 2012 for examples of tables for selecting statistical tests). Although an appropriate test may be identified using these tables, this process of test selection is broadly a-theoretical and does not help students to understand the following questions:

- Why may a particular test be applied?
- What assumptions are / have been made?
- How can problems be tested and rectified?
- How can the test be generalized to accommodate different designs, different types of data or additional variables?

or even

- How should the data be structured to enable a statistical analysis to be carried out?

Many of these difficulties can, however, be addressed by the use of regression models. These allow statistical tests to be easily identified, detailed diagnostics to be applied, additional variables to be added and different research designs accommodated. The use of regression models is particularly relevant as many traditional tests are simply standard regression models that are restricted to certain conditions[1]. These restrictions were once necessary, since the tests were developed at a time when the complexity of arithmetical calculations needed for a more unified approach excelled the resources generally available. The widespread availability of microcomputers and statistical software, such as the R language (R Development Core Team, 2012), have changed this situation completely. Pencils, paper, books of log-tables and slide-rules have been replaced by efficient, easy to use statistical computing environments that have facilitated the migration from a plethora of special cases to an approach based on a single, logically-consistent structure.

Regression analyses not only provide identical results to the traditional parametric tests, they also provide a number of additional benefits. One of the major benefits of using GLMs is that the regression models have a consistent underlying theory that can be extended to model different types of data, including ordered and unordered categorical and count data. This enables the large range of tests available for parametric and non-parametric data to be replaced with a small number of generalized linear models (GLMs) that provides the user with a unified theory for test selection and also offers significant advantages with respect to the quality of analysis.

Put concisely, the continued propagation of ‘pen and paper techniques’ represents a bounding factor for both the modern organization, be it educational or commercial, and for the individual, who chooses to remain in the statistical Stone Age. In their excellent introduction to statistical thinking, Diggle and Chetwynd summarize these thoughts very elegantly: “...restricting the statistical recipe book to explicit formulae is unnecessary. Far better [...] to adopt a principled approach to statistical analysis, in which the old ‘which test can I use on my data?’ mantra is replaced by ‘how can I use statistical thinking to help me get the best possible answer to my scientific question?’ ”

1 That is, identical results to those obtained using the traditional tests can be obtained using regression. As regression is far more flexible, the traditonal tests are, essentially, redundant.

A short comment on statistical models.

Although the notion of a statistical model will be familiar to the reader on an intuitive level, the concept is used almost universally without even the most basic explanation. This tutorial advocates statistical modelling within the uniform, overarching framework of generalized linear models. It is therefore perhaps wise to outline the concept somewhat, not only to avoid misconceptions but also to motivate the reader to mull over statistical goals, rather than unconsciously using the 'model' as just another unavoidable ingredient in the statistical soup.

It would also be all too easy to start an abstract mathematical discussion of statistical models, filling page after page with notation that may be logically satisfying, but probably bereft of any useful content for the task at hand. The interested reader is directed towards Dekking et al. (2005) for an excellent formal introduction to statistical thought, and to McCullagh (2002), which offers an exacting but interesting discussion of some aspects of statistical models.

Lindsey (1998) fixes the role of statistical modelling firmly within the context of the subject of study:

> Any data collected contain a mass of information. The problem is for you to extract that part of it that is relevant to the questions to be answered by your study, in the simplest and most understandable way possible. This essentially involves checking for pertinent patterns and anomalies in the data. This is a basic role of statistical models: to simplify reality in a reasonable and useful way, a way that you can empirically check with the data.

"The formalization of variability, as an approximation to reality, is known as a statistical model." (Lindsey, 1995). This formalization of variability (on the sample space of interest) occurs through one or more (parametrized) statistical distributions, producing a mathematical abstraction of reality that captures the fundamental characteristics of interest.

The importance of the connection between the model and reality cannot be stressed too strongly. Concentrating on goodness of fit attributes between data and model, while condemning such thoughts as "does this really make sense?" to a subservient role may cause little real damage in academic circles, but when perpetrated in the real world may have grave consequences. We advocate the use of a consistent, understandable and flexible framework in order to free up practitioners so that they may concentrate on the subject matter, and not to ease the manufacture of sub-standard analyses:

> If a particular model (parametrization) does not make biological sense, this is a reason to exclude it from the set of candidate models, particularly in the case where causation is of interest. (Burnham and Anderson, 2002)

A regression-based approach.

This section discusses how traditional tests can be mapped onto linear models. In its simplest form, the relationship between a dependent variable and an independent variable (also known as the response and explanatory variables) can be represented by the model

Y ~ X

The ~ symbol signifies a link between the variables X and Y, and the form of this relationship depends on the distribution of the Y variable (see McCullagh and Nelder, 1989). This model simply states that variable 'Y' may be related to, i.e. dependent on, variable 'X'. Examples of such relationships are:

- The score achieved in a mathematics test may be related to gender.
- The attitude of an employee may be related to their status within the company.
- The likelihood of injury may be related to the wearing of safety equipment.

Y can be measured on any scale (see Stevens, 1946 for a discussion of measurement scales) and helps identify the specific technique used for the model. If, for example, Y is continuous, an ordinary least squares (OLS) regression model may be used; should Y be categorical, a logit regression model may be more appropriate, and if Y is a count or frequency variable, a Poisson regression may be used. With the application of appropriate dummy variable coding techniques, variable X can also be measured on any scale (see Hutcheson, 2011b, for a tutorial on dummy-variable coding).

When X is an unordered categorical variable (e.g. gender, group, or company), the model Y ~ X compares scores between the group categories for independent observations, i.e., different subjects provide scores for each category. This model is an independent groups design and is represented using a regression model format as

Y ~ group

This model represents the traditional group tests for independent samples. For example, when Y is a continuous variable, the model is an OLS regression which is equivalent to an independent groups ANOVA (or a t-test, if there are just two group categories). If Y is ordered categorical, a proportional-odds logit model may be used. This replaces non-parametric tests such as the Mann-Whitney and Kruskal-Wallis. If Y is unordered categorical, a multinomial logit model may be used which is equivalent to a chi-square contingency table analysis (contingency table cell counts can also be analyzed using a Poisson regression model that directly models count data[2]). Many of these analyses and the "equivalent" traditional tests are demonstrated below.

The regression models are particularly versatile as additional information available to the analyst can be accounted for by adding extra variables into the model. For example, if the samples from the study are drawn from different schools or companies, this information may be simply added into the model

2 The link (~) for count data is the log – which equates to a log-linear model.

Y ~ group + company

Extending the simple model Y ~ X, the above states that variable ‘Y’ may be related to, i.e. dependent on, variables ‘group’ and ‘company’. The regression model is, essentially, the same as before and a regression technique based on the measurement of Y is applied. It is not easy, however, to adapt the traditional tests to account for additional variables, as this changes the test that may be applied.

The ability to add additional variables has an added benefit in that it makes it easy to account for different research designs. For example, matched-groups designs can be analysed by simply adding the information about matching. For example, in a repeated measures design, a single subject is represented in more than one group category (for example, when individual subjects rate a number of different stores). A regression model for a matched-groups design simply includes both the 'group' and 'subject' variables...

Y ~ group + subject

The additional information about subject is added to the model as an unordered categorical variable. This model represents the traditional group tests for dependent samples. For example, when Y is a continuous variable, the model is an OLS regression, which is equivalent to an dependent groups ANOVA (or a paired t-test, if there are just two group categories). If Y is ordered categorical, a proportional-odds logit model may be used, which replaces the non-parametric tests such as the Wilcoxon and Friedman. If Y is unordered categorical, a multinomial logit model may be used which is equivalent to a multi-way contingency table analysis.

The important thing to note here is that all of the traditional group group tests can be analysed using the generalized linear model ‘Y ~ group’ and an appropriate dummy coding scheme.[3] In addition to providing a simpler structure for group-tests, the regression models provide a number of other advantages. The ability to add additional variables to the model enables more complex hypotheses to be tested (these models reproduce or replace the two-way ANOVAs, ANCOVAs and the equivalent non-parametric tests). This is particularly important given the non-experimental nature of many research studies.[4] The regression models also allow easy access to diagnostic tools and data transformations.

These models may appear simplistic, but they reproduce or replace the traditional hypothesis tests and provide a much simpler and more powerful system of analysis. Table 1 shows a list of traditional tests[5] (ordered according to the number of independent variables and the research design) and the corresponding, i.e. replacement, regression models.

---

3 The application of dummy coding should not present any difficulty, as modern statistical packages automatically dummy code categorical variables (see, for example, R development core team, 2012).

4 Potentially confounding variables are controlled in the experimental design (random selection of subjects and random assignment to groups etc.). Although such designs are assumed by traditional tests, they are rarely used in management research (most studies being survey-based).

5 These are just a few that are available.

Table 1. Traditional test and their equivalent regression models.

Traditional test	GLM model
1 independent variable (independent groups)	
t-test (unrelated)	score ~ group
Mann-Whitney	
1-way ANOVA (unrelated)	
Kruskal-Wallis	
Jonck-heere Trend	
Chi-square (contingency table)	
1 independent variable (dependent groups)	
t-test (related)	score ~ group + subject
Wilcoxon	
1-way ANOVA (related)	
Friedman	
Page's L trend	
2 or more independent variables (independent groups)	
Complex selection of multi-way ANOVA and ANCOVA tests for unrelated designs containing additional variables...	score ~ group + additional
2 or more independent variables (dependent groups)	
Complex selection of multi-way ANOVA and ANCOVA tests available for related designs containing additional variables...	score ~ group + additional + subject

Demonstrating the analyses.

It is useful to demonstrate that the traditional parametric tests can be reproduced and that the non-parametric tests may be replaced with regression models. Table 2 shows an example dataset which contains minimal data required to run the analyses. It should be noted that analyses are just for demonstration and are not intended to be accurate models (the dataset is far too small to represent the population). All statistics have been computed in R (R Development Core Team, 2012) with analysis of deviance tables reported for all regression models (see Fox and Weisberg, 2011 and Hutcheson and Moutinho, 2008, for discussions of deviance in regression).

Table 2: An example dataset for illustrating traditional tests and regression models. These data can be downloaded from www.research-training.net/TESTselectionDATA.csv (the data are saved in a coma-separated-variable format).

score.num	**score.ord**	**group**	**subject**	**age**	**region**
63	B	A	subj01	18	south
67	B	A	subj02	63	West
65	B	A	subj03	25	north
59	C	A	subj04	43	north
55	C	A	subj05	21	south
71	A	B	subj01	17	north
69	B	B	subj02	24	West
70	A	B	subj03	27	south
75	A	B	subj04	19	south
59	C	B	subj05	28	West
42	D	C	subj01	72	north
48	D	C	subj02	59	north
51	C	C	subj03	49	south
41	B	C	subj04	58	West
55	C	C	subj05	45	north

A. Reproducing traditional parametric tests using regression.

Traditional tests for parametric data (for example, t-tests, ANOVA, ANCOVA) are special cases of regression and can be reproduced using OLS regression models. The following examples show the same analyses conducted on 2 and 3 category groups on data collected from independent and dependent group studies.

Independent groups:

A one-way ANOVA or an OLS regression model may be used to test whether there is a difference in 'score.num' (a continuous variable) between the groups (indicated by variable 'group') for data collected using an independent-groups design.

Traditional test: one way ANOVA

A traditional one way ANOVA results in the following output...

```
             Df   Sum Sq   Mean Sq   F value     Pr(>F)
group         2   1190.5     595.3     19.06   0.000188
Residuals    12    374.8      31.2

    mean         sd   data:n
A   61.8   4.816638        5
B   68.8   5.932959        5
C   47.4   5.941380        5
```

Regression model: score.num ~ group

A regression model of score.num using group as an explanatory variable, results in the following output...

```
glm(formula = score.num ~ group, family = gaussian(identity))

Coefficients:
              Estimate   Std. Error   t value   Pr(>|t|)
(Intercept)     61.800        2.499    24.727   1.16e-11
group[T.B]       7.000        3.535     1.980    0.07105
group[T.C]     -14.400        3.535    -4.074    0.00154

Analysis of Deviance Table (Type II tests)
Response: score.num
                SS   Df        F      Pr(>F)
group       1190.5    2   19.059   0.0001884
Residuals    374.8   12
```

The one way ANOVA and the OLS regression model provide identical outputs. Both tests show an F-value of 19.06 and also show the differences between the groups. The intercept term for the regression model shows the mean value for the reference category (group A, 61.8), and the individual parameters show how different each group is from the reference; group B is 7 units higher (i.e. has a value of 68.8) and group C is 14.4 units lower (i.e. has a value of 47.4)[6].

6 Using regression, different dummy coding techniques can be used to provide different group comparisons. For example, a different reference category can be chosen, or each group can be compared to the mean of all groups (see

Dependent-groups:

In order to demonstrate the analysis of data collected using a repeated measures design (and a different number of groups), the following analysis compares just two of the groups reported in Table 2 – groups A and C. These data can be analysed using a paired t-test, or, equivalently, an OLS regression model.

Traditional test: Paired t-test

In order to run a traditional paired t-test, the data need to be presented using a different structure (see Hutcheson, 2011a, for a discussion of data structures).

scoreA	scoreC
63	42
67	48
65	51
59	41
55	55

A traditional paired t-test results in the following output...

```
Paired t-test

data:  score.numA and score.numC
t = 3.8133,  df = 4,  p-value = 0.01889
mean of the differences
                   14.4
```

Regression model: score.num ~ group + subject

A regression model of a repeated measures design is identical to the dependent groups design, except that information about the subject is added to the model. This model requires each variable to be represented as a single column of data. The structure of the data in Table 2 is therefore appropriate and does not need to be changed.

An OLS regression model of score.num using group and subject as explanatory variables results in the following output...

Hutcheson, 2011c, for a discussion of dummy coding techniques).

```
Coefficients:
                    Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)           59.700       4.625   12.908  0.000208
group[T.C]           -14.400       3.776   -3.813  0.018886
subject[T.subj02]      5.000       5.971    0.837  0.449477
subject[T.subj03]      5.500       5.971    0.921  0.409098
subject[T.subj04]     -2.500       5.971   -0.419  0.696935
subject[T.subj05]      2.500       5.971    0.419  0.696935

Analysis of Deviance Table (Type II tests)
Response: score
              SS   Df        F    Pr(>F)
group      518.4    1   14.541   0.01889
subject     91.4    4    0.641   0.66148
```

Although the results from the two analyses may look quite different, they are identical. Both techniques show the same significance for the group. The paired t-test value of 3.8133 is identical to the regression F-ratio of 14.541 (the square of the t-value is equivalent to the F-ratio; $3.8133^2 = 14.541$) and both show the difference between the groups as being 14.4 units. The regression analysis, however, provides more information (for example, the significance of the subject variable) and also allows the option of including additional variables and computing a range of diagnostics.

B. Replacing the traditional non-parametric tests using regression.

When the response variable is not continuous (or does not have a Gaussian distribution), a traditional non-parametric test can be used. There are a huge number of tests available (see, for example, Siegal and Castellan, 1988) which can cause confusion for test selection. Many of the standard non-parametric tests can, however, be replaced with regression models (in this case, the proportional-odds model). The regression models reported here do not reproduce the non-parametric test results – they replace them (the statistics and significance levels are different). The following examples show the same analyses conducted on multi-category groups on ordered data collected from independent and dependent group studies.

Independent-groups:

A Kruskal-Wallis rank sum test or a regression model may be used to test whether there is a difference in ‘score.ord’ (an ordered categorical variable) between the groups (indicated by variable ‘group’) for data collected using an independent-groups design.

Traditional test: Kruskal-Wallis rank sum test

A traditional Kruskal-Wallis rank sum test results in the following output...

```
data:  score.ord by group
Kruskal-Wallis chi-squared = 6.285, df = 2, p-value = 0.04317
```

Regression model: score.ord ~ group

A regression model of score.num using group as an explanatory variable results in the following output...

```
Coefficients:
             Value Std. Error t value
group[T.B] -2.285      1.419  -1.610
group[T.C]  2.120      1.378   1.538

Intercepts:
    Value   Std. Error t value
A|B -2.1299  1.1575    -1.8402
B|C  0.2448  0.8075     0.3031
C|D  2.7243  1.2743     2.1379


Analysis of Deviance Table (Type II tests)
Response: score.ord
      LR Chisq Df Pr(>Chisq)
group   8.9694  2    0.01128 *
```

The traditional test shows a significant difference between the groups and a p-value of 0.04317. This is different to that provided by the regression model, which shows a p-value of 0.01128. Given the very small sample size, it is not appropriate to interpret these values with respect to the population. The regression model does, however, offer a number of advantages over the Kruskal-Wallis including the ability to test ordinality (the values of the intercepts provide clues about this) and the appropriateness of using an ordered regression model (this can be checked using the proportional-odds assumption test), the ability to add additional variables to the model and the availability of diagnostics.

Dependent groups:

In order to demonstrate the analysis of data collected using repeated measures designs, the following analysis compares the ordered score data (variable 'score.ord') for the groups. These data can be analysed using a Friedman rank sum test, or a regression model.

Traditional test: Friedman rank sum test

In order to run a Friedman rank sum test, the data need to be presented using a structure similar to that used above for the paired t-test. A traditional Friedman rank sum test results in the following output...

```
Friedman chi-squared = 5.7333, df = 2, p-value = 0.05689
```

Regression model: score.ord ~ group + subject

A regression model of a repeated measures design is identical to the regression model used for the dependent groups design, except that information about the subject is added to the model. This model requires each variable to be represented as a single column of data. The structure of the data in Table 2 is therefore appropriate and does not need to be changed. A regression model of score.ord using group and subject as explanatory variables results in the following output...

```
Coefficients:
                   Value Std. Error t value
group[T.B]        -2.609      1.480 -1.7626
group[T.C]         2.674      1.545  1.7308
subject[T.subj02]  1.171      1.629  0.7187
subject[T.subj03] -1.120      1.676 -0.6685
subject[T.subj04] -1.314      1.905 -0.6896
subject[T.subj05]  2.136      1.782  1.1988

Intercepts:
    Value   Std. Error t value
A|B -2.4520  1.6166    -1.5168
B|C  0.5433  1.4070     0.3862
C|D  3.7669  1.8582     2.0271


Analysis of Deviance Table (Type II tests)
Response: score.ord
        LR Chisq Df Pr(>Chisq)
group    11.6977  2   0.002883
subject   5.2682  4   0.260873
```

The Friedman test shows a difference between the groups equivalent to a p-value of 0.057. This is different to that provided by the regression model, which shows a p-value of 0.003. Given the very small sample size, it is not appropriate, however, to interpret these values with respect to the population. The regression model offers a number of advantages over the Friedman including the ability to test ordinality (the values of the intercepts provide clues about this) and the appropriateness of using an ordered regression model (this can be checked using the proportional-odds assumption test), the ability to add additional variables to the model and the availability of diagnostics.

C. Analysing categorical data using regression.

The relationship between categorical variables is traditionally analyzed using tests based on chi-square that are applied to contingency tables. The statistics from these tests can, however, also be reproduced using generalized regression models which offer significant advantages over the traditional tests (for example, the ability to add additional variables into the model and the application of the same unified underlying theory for data analysis). A traditional contingency table analysis is compared to regression models using the data shown in Table 2.

Traditional test:

A contingency table analysis may be used to investigate the relationship between 'group' and 'region'. The contingency table derived from the data in Table 2 is

		region		
		north	south	west
	A	2	2	1
group	B	1	2	2
	C	3	1	1

A traditional contingency-table test results in the following output...

```
        Pearson's Chi-squared test

X-squared = 1.9,    df = 4,      p-value = 0.7541
G-squared = 1.955,  df = 4,      p-value = 0.7441
```

Regression models:

There are two regression models that can be used to replicate the contingency table analysis provided above. Cell-count can be modelled using a GLM with a Poisson link (a GLM applied to count data), or one of the variables can be modelled using a GLM with a logit link (a GLM applied to unordered categorical data). Both of these are shown below.

A model of cell-count:

The model used here predicts the cell count using 'group' and 'region' as explanatories. The regression model is...

cell-count ~ group * region

where the term 'group * region' includes the main effects and the interaction (i.e, group + region + group:region). As we wish to assess the relationship between group and region, it is the interaction term between these that we are interested in. As the response variable is a count, we use a Poisson model which simply uses a Poisson link between the response and explanatory variables. The model we run is therefore very similar to previous models except that a Poisson link is used. A regression model of cell-count using group and region as explanatory variables results in the following output...

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	6.931e-01	7.071e-01	0.980	0.327
group[T.B]	-6.931e-01	1.225e+00	-0.566	0.571
group[T.C]	4.055e-01	9.129e-01	0.444	0.657
region[T.s]	-4.254e-16	1.000e+00	0.000	1.000
region[T.w]	-6.931e-01	1.225e+00	-0.566	0.571
group[T.B]:region[T.s]	6.931e-01	1.581e+00	0.438	0.661
group[T.C]:region[T.s]	-1.099e+00	1.528e+00	-0.719	0.472
group[T.B]:region[T.w]	1.386e+00	1.732e+00	0.800	0.423
group[T.C]:region[T.w]	-4.055e-01	1.683e+00	-0.241	0.810

Analysis of Deviance Table (Type II tests)

Response: count

	LR Chisq	Df	Pr(>Chisq)
group	0.00000	2	1.0000
region	0.40271	2	0.8176
group:region	1.95455	4	0.7441

The interaction term in the Poisson model is identical to the G-squared statistic reported as part of a standard contingency table analysis. The Poisson regression model, however, provides a lot more information (for example, the comparisons between individual cells) and also allows the model to be extended with additional variables.

A model of an unordered categorical variable:

It is also possible to replicate the contingency table analysis by directly modelling one of the unordered categorical variables. A GLM of the variable ‘group’ is simply...

group ~ region

As the response variable is unordered categorical, we use the multinomial technique which automatically selects a logit link between the response and explanatory variables. The model we run is therefore very similar to previous models except that a logit link is used via the multinomial technique. A regression model of group using region as an explanatory variable results in the following output...

```
Coefficients:
  (Intercept) region[T.south] region[T.west]
B  -0.6931430       0.6931407      1.3862642
C   0.4054576      -1.0985788     -0.4054736

Std. Errors:
  (Intercept) region[T.south] region[T.west]
B   1.2247413        1.581138       1.732045
C   0.9128709        1.527519       1.683245

Value/SE (Wald statistics):
  (Intercept) region[T.south] region[T.west]
B  -0.5659505       0.4383810      0.8003628
C   0.4441566      -0.7191917     -0.2408880


Analysis of Deviance Table (Type II tests)

Response: group
       LR Chisq Df Pr(>Chisq)
region   1.9546  4     0.7441
```

The multinomial logit model 'group ~ region' provides the same statistics as the Poisson model above (chi-sqaure = 1.9546, which is also the same as that for the contingency table analysis). The multinomial model, however, offers a number of advantages, such as more detailed information about the relations between categories and the ability to add additional variables into the model (including variables measured on a continuous scale).

Conclusion

This tutorial has demonstrated that the traditional models for analysing grouped data may be represented or replaced by GLMs. The many seemingly different tests that cause so many difficulties for those learning statistics may be replaced by a single model...

$$Y \sim X_1 + \ldots + X_n$$

Models of continuous, count, ordered and unordered categorical data, are all based on this same basic model. This model can be easily adapted to account for experimental design and also allows sophisticated diagnostics to be applied. The regression model also defines the variables needed to enter into the model and thereby provides the structure of the data, which is a difficulty commonly encountered by students. One of the biggest advantages to using GLMs, however, is the common underlying theory and the ease with which this can be taught.

Given these advantages, there seems little reason not to base statistical analysis courses on GLMs. It is time that test selection and teaching moved into the 21st century.

Graeme D Hutcheson (Manchester University, UK)
Les Schaefer (Independent Consultant, Hamburg, Germany)

References/Bibliography

Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference*. Springer.

Dekking, F.M., Kraaikamp, C., Lopuhaä, H.P., and Meester, L.E. (2005). *A Modern Introduction to Probability and Statistics: Understanding Why and How*. Springer.

Diggle, P.J. and Chetwynd, A.G. (2011). *Statistics and Scientific Method*. Oxford University Press.

Fox, J. and Weisberg, S. (2011). *An R companion to applied regression (2nd edition).* Sage Publications.

Hutcheson, G. D. (2011a). *Data set Structure*. In The Sage Dictionary of Quantitative Management Research by Moutinho, L. and Hutcheson, G. D. (editors), pp: 74-77. Sage Publications.

Hutcheson, G. D. (2011b). Categorical Explanatory Variables. *Journal of Modelling in Management*, **6,2**:225-236.

Hutcheson, G. D. (2011c). *Dummy variable coding.* In The Sage Dictionary of Quantitative Management Research by Moutinho, L. and Hutcheson, G. D. (editors), pp: 98-102. Sage Publications.

Hutcheson, G. D. and Moutinho, L. (2008). *Statistical Modeling for Management*. Sage Publications.

Lindsey, J.K. (1995). *Introductory Statistics: A Modelling Approach*. Oxford University Press.

Lindsey, J.K.. (1998). *Revealing Statistical Principles.* Available: www.commanster.eu/ms/principles.pdf. Last accessed 10. August 2012.

McCullagh, P. (2002). What is a statistical model?. *The Annals of Statistics*. **30**,**5**:1225-1310.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*. Chapman and Hall.

Motulsky, H. (2010). *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking (*2nd edition). Oxford University Press.

Motulsky, H.. (2012). *Frequently Asked Questions: Choosing a statistical test.* Available: http://www.graphpad.com/support/faqid/1790/. Last accessed 10th August 2012.

R Development Core Team. (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*. **133**,**2684**:677-680.