

Chapter 1 : Discovering Statistics Using IBM SPSS Statistics : Andy Field :

"In this brilliant new edition Andy Field has introduced important new introductory material on statistics that the student will need and was missing at least in the first edition. This book is the best blend that I know of a textbook in statistics and a manual on SPSS.

Andy Field 3rd Edition, Chapter 1 Why is my evil lecturer forcing me to learn statistics? A categorical variable is made up of categories. A categorical variable that you should be familiar with already is your species e. You are a human or a cat or a fruit bat: A categorical variable is one that names distinct entities. In its simplest form it names just two distinct types of things, for example male or female. This is known as a binary variable. In all cases there are just two categories and an entity can be placed into only one of the two categories. When two things that are equivalent in some sense are given the same name or number, but there are more than two possibilities, the variable is said to be a nominal variable. It should be obvious that if the variable is made up of names it is pointless to do arithmetic on them if you multiply a human by a cat, you do not get a hat. However, sometimes numbers are used to denote categories. For example, the numbers worn by players in a rugby or football soccer team. In rugby, the numbers of shirts denote specific field positions, so the number 10 is always worn by the fly-half e. These numbers do not tell us anything other than what position the player plays. We could equally have shirts with FH and H instead of 10 and 1. A number 10 player is not necessarily better than a number 1 most managers would not want their fly-half stuck in the front of the scrum! It is equally as daft to try to do arithmetic with nominal scales where the categories are denoted by numbers: The only way that nominal data can be used is to consider frequencies. For example, we could look at how frequently number 10s score tries compared to number 4s. So far the categorical variables we have considered have been unordered e. When categories are ordered, the variable is known as an ordinal variable. Ordinal data tell us not only that things have occurred, but also the order in which they occurred. However, these data tell us nothing about the differences between values. Imagine we went to a beauty pageant in which the three winners were Billie, Freema and Elizabeth. These categories are ordered. In using ordered categories we now know that the woman who won was better than the women who came second and third. We still know nothing about the differences between categories, though. Billie might have been an easy victor, getting much higher ratings from the judges than Freema and Elizabeth, or it might have been a very close contest that she won by only a point. Ordinal data, therefore, tell us more than nominal data they tell us the order in which things happened but they still do not tell us about the differences between points on a scale. The next level of measurement moves us away from categorical variables and into continuous variables. A continuous variable is one that gives us a score for each person and can take on any value on the measurement scale that we are using. The first type of continuous variable that you might encounter is an interval variable. Interval data are considerably more useful than ordinal data and most of the statistical tests in this book rely on having data measured at this level. To say that data are interval, we must be certain that equal intervals on the scale represent equal differences in the property being measured. For this scale to be interval it must be the case that the difference between helpfulness ratings of 1 and 2 is the same as the difference between say 3 and 4, or 4 and 5. Similarly, the difference in helpfulness between ratings of 1 and 3 should be identical to the difference between ratings of 3 and 5. Variables like this that look interval and are treated as interval are often ordinal " see Jane Superbrain Box 1. This is known as a frequency distribution, or histogram, which is a graph plotting values of observations on the horizontal axis, with a bar showing how many times each value occurred in the data set. Frequency distributions can be very useful for assessing properties of the distribution of scores. We will find out how to create these types of charts in Chapter 4. Frequency distributions come in many different shapes and sizes. It is quite important, therefore, to have some general descriptions for common types of distributions. In an ideal world our data would be distributed symmetrically around the centre of all scores. As such, if we drew a vertical line through the centre of the distribution then it should look the same on both sides. This is known as a normal distribution and is characterized by the bell-shaped curve with which you might already be familiar. This shape basically implies that the majority of scores lie around

the centre of the distribution so the largest bars on the histogram are all around the central value. Also, as we get further away from the centre the bars get smaller, implying that as scores start to deviate from the centre their frequency is decreasing. As we move still further away from the centre our scores become very infrequent the bars are very short. Many naturally occurring things have this shape of distribution. For example, most men in the UK are about 175 cm tall; [12] some are a bit taller or shorter but most cluster around this value. There will be very few men who are really tall or really short. An example of a normal distribution is shown in Figure 1. There are two main ways in which a distribution can deviate from normal: Skewed distributions are not symmetrical and instead the most frequent scores the tall bars on the graph are clustered at one end of the scale. So, the typical pattern is a cluster of frequent scores at one end of the scale and the frequency of scores tailing off towards the other end of the scale. A skewed distribution can be either positively skewed the frequent scores are clustered at the lower end and the tail points towards the higher or more positive scores or negatively skewed the frequent scores are clustered at the higher end and the tail points towards the lower or more negative scores. Distributions also vary in their kurtosis. Kurtosis, despite sounding like some kind of exotic disease, refers to the degree to which scores cluster at the ends of the distribution known as the tails and how pointy a distribution is but there are other factors that can affect how pointy the distribution looks – see Jane Superbrain Box 2. A distribution with positive kurtosis has many scores in the tails a so-called heavy-tailed distribution and is pointy. This is known as a leptokurtic distribution. In contrast, a distribution with negative kurtosis is relatively thin in the tails has light tails and tends to be flatter than normal. This distribution is called platykurtic. Ideally, we want our data to be normally distributed. In a normal distribution the values of skew and kurtosis are 0. If a distribution has values of skew or kurtosis above or below 0 then this indicates a deviation from normal: There is a standard formation in rugby and so each team has 15 specific positions that must be held constant for the game to be played. When the first player arrives, you have the choice of 15 positions in which to place this player. You place his name in one of the slots and allocate him to a position. When the next player arrives, you have the choice of 14 positions but you still have the freedom to choose which position this player is allocated. However, as more players arrive, you will reach the point at which 14 positions have been filled and the final player arrives. With this player you have no freedom to choose where they play – there is only one position left. Therefore there are 14 degrees of freedom; that is, for 14 players you have some degree of choice over where they play, but for 1 player you have no choice. The degrees of freedom is one less than the number of players. In statistical terms the degrees of freedom relate to the number of observations that are free to vary. If we take a sample of four observations from a population, then these four scores are free to vary in any way they can be any value. Thus we hold one parameter constant. Say that the mean of the sample was 10; then we assume that the population mean is 10 also and we keep this value constant. With this parameter fixed, can all four scores from our sample vary? The answer is no, because to keep the mean constant only three values are free to vary. Therefore, if we hold one parameter constant then the degrees of freedom must be one less than the sample size. Going beyond the data 2. This is where we use the standard error. Many students get confused about the difference between the standard deviation and the standard error usually because the difference is never explained clearly. We have already learnt that social scientists use samples as a way of estimating the behaviour in a population. Imagine that we were interested in the ratings of all lecturers so, lecturers in general were the population. We could take a sample from this population. When someone takes a sample from a population, they are taking one of many possible samples. If we were to take several samples from the same population, then each sample has its own mean, and some of these sample means will be different. For each of these samples we can calculate the average, or sample mean. This illustrates sampling variation: We can actually plot the sample means as a frequency distribution, or histogram, [4] just like I have done in the diagram. This distribution shows that there were three samples that had a mean of 3, means of 2 and 4 occurred in two samples each, and means of 1 and 5 occurred in only one sample each. The end result is a nice symmetrical distribution known as a sampling distribution. A sampling distribution is simply the frequency distribution of sample means from the same population. So how do we determine the accuracy of the population mean? Think back to the discussion of the standard deviation. We used the standard deviation as a

measure of how representative the mean was of the observed data. Small standard deviations represented a scenario in which most data points were close to the mean, a large standard deviation represented a situation in which data points were widely spread from the mean.

Chapter 2 : Discovering Statistics Using SPSS | eBay

With an exciting new look, math diagnostic tool, and a research roadmap to navigate projects, this new edition of Andy Field's award-winning text offers a unique combination of humor and step-by-step instruction to make learning statistics compelling and accessible to even the most anxious of students.

Chapter 3 : Discovering Statistics Using IBM SPSS Statistics â€” Discovering Statistics

Discovering Statistics Using SPSS is an amazing tool that I discovered in It is very engaging - using simple and straightforward language for complex statistical concepts. It loses a star as I find it less useful as a reference for regressions which I think is a shame.

Chapter 4 : Discovering Statistics Using SPSS, 3rd Edition, by Andy Field | Meng Hu's Blog

Welcome to Discovering calendrierdelascience.com website brings together many of my free resources to help you to get to grips with statistics. These include handouts, tutorials, extracts from my books, videos and blogs.

Chapter 5 : Discovering Statistics Using SPSS - Andy Field - Google Books

4th edition andy field statistics using ibm spss statistics discovering and sex and drugs and rock 'n' roll field calendrierdelascience.com 3 12/11/ pm.

Chapter 6 : calendrierdelascience.com | Discovering Statistics Using IBM SPSS | | Andy Field | Boeken

Discovering statistics using SPSS Andy Field First edition published Second edition published Apart from any fair dealing for the purposes of.

Chapter 7 : Discovering Statistics Using SPSS by Andy Field

Andy Field's bestselling Discovering Statistics Using SPSS 4th Edition, already an immensely comprehensive textbook - taking students from first principles to advanced statistical concepts, and all the while grounding knowledge through the use of SPSS - now focuses on providing essential updates, better accessibility to its key features, more.

Chapter 8 : Discovering Statistics Using IBM SPSS Statistics (ebook) by Andy Field |

Find all the study resources for Discovering Statistics Using SPSS by Andy P. Field.

Chapter 9 : Discovering Statistics â€” The adventure begins â€”

The fourth edition of Discovering Statistics will transform students who approach statistics with fear and loathing into adroit statistics users who understand key statistical concepts. Field's book is a practical 'how to' guide for conducting and understanding basic and advanced statistical analyses using IBM SPSS Statistics.