It was remarked earlier that the discreteness of the distribution of r, the number of 'successes', made inferences from proportions a little different from those based on a variable with a continuous distribution and we now discuss these differences. For a continuous variable an exact significance test would give the result P < 0 05 for exactly 5% of random samples drawn from a population in which the null hypothesis were true, and a 95% confidence interval would contain the population value of the estimated parameter for exactly 95% of random samples. Neither of these properties is generally true for a discrete variable. Consider a binomial variable from a distribution with n = 10 and π = 0 5 (Table 2.4, p. 66). Using the exact test, for the hypothesis that π = 0 5, significance at the 5% level is found only for r = 0, 1, 9 or 10 and the probability of one or other of these values is 0.022. Therefore, a result significant at the 5% level would be found in only 2.2% of random samples if the null hypothesis were true. This causes no difficulty if the precise level of P is stated. Thus if r = 1 we have that P = 0.022, and a result significant at a level of 0.022 or less would occur in exactly 2.2% of random samples. The normal approximation with continuity correction is then the best approximate test, giving, in this case, P =0.027.

A similar situation arises with the confidence interval. The exact confidence limits for the binomial parameter are conservative in the sense that the probability of including the true value is *at least* as great as the nominal confidence coefficient. This fact arises from the debatable decision to include the observed value in the calculation of tail-area probabilities. The limits are termed **'exact'** because they are obtained from exact calculations of the binomial distribution, rather than from an approximation, but not because the confidence coefficient is achieved exactly. This problem cannot be resolved, in the same way as for the significance test, by changing the confidence coefficient. First, this is difficult to do but, secondly and more importantly, whilst for a significance test it is desirable to estimate P as precisely as possible, in the confidence interval approach it is perfectly reasonable to specify the confidence coefficient in advance at some conventional value, such as 95%. The approximate limits using the continuity correction also tend to be conservative. The limits obtained by methods 2 and 3, however, which ignore the continuity correction, will tend to have a probability of inclusion nearer to the nominal value. This suggests that the neglect of the continuity correction is not a serious matter, and may, indeed, be an advantage.

The problems discussed above, due to the discreteness of the distribution, have caused much controversy in the statistical literature, particularly with the analysis of data collected to compare two proportions, to be discussed in §4.9. One approach, suggested by Lancaster (1952, 1961), is to use **mid-P** values, and this approach has been advocated more widely recently (Williams, 1988; Barnard. 1989; Hirji, 1991; Upton, 1992). The mid-P value for a one-sided test is obtained by including in the tail only one-half of the probability of the observed sample. Thus for a binomial sample with r observed out of n where r > nπ, the one-sided mid-P value testing the hypothesis that π = π_0 will be

$$\text{mid-P} = 0.5P[r] + P[r+1] + ... + P[n]$$

It has to be noted that the mid-P value is not the probability of obtaining a significant result by chance when the null hypothesis is true. Again,

consider a binomial variable from a distribution with n = 10 and   = 0 5 (Table 2.4, p. 66). For the hypothesis that   =0.5, a mid-P value less than 0.05 would be found only for r = 0, 1, 9 or 10, since the mid-P value for r = 2 is 2[0 0010 + 0 0098 +0.5(0 0439)] = 0 0655, and the probability of one or other of these values is 0022.   Barnard (1989) has recommended quoting both the P and the mid-P values, on the basis that the former is a measure of the statistical significance when the data under analysis are judged alone, whereas the latter is the appropriate measure of the strength of evidence against the hypothesis under test to be used in combination with evidence from other studies. This arises because the mid-P value has the desirable feature that, when the null hypothesis is true, its average value is 0.5 and this property makes it particularly suitable as a measure to be used when combining results from several studies in making an overall assessment (meta-analysis; Chapter 7). Since it is rare that the results of a single study are used without support from other studies, our recommendation is also to give both the P and mid-P values, but to give more emphasis to the latter.

Corresponding to mid-P values are mid-P confidence limits, calculated as those values which, if taken as the null hypothesis value, give a corresponding mid-P value, that is, the 95% limits correspond to one-sided mid-P values of 0-025.

Where a normal approximation is adequate, P values and mid-P values correspond to test statistics calculated with and without the correction for continuity respectively. Correspondingly, confidence intervals and mid-P confidence intervals can be based on normal approximations, using and

ignoring the continuity correction respectively. Thus the mid-P confidence limits for a binomial probability would be obtained using method 2 rather than method 1 (p. 121).   Where normal approximations are inadequate, the mid-P values are calculated by summing the appropriate probabilities. The mid-P limits are more tedious to calculate, as they are not included in standard sets of tables and there is no direct formula corresponding to (4.10). The limits may be obtained fairly readily using a personal computer or programmable calculator by setting up the expression to be evaluated using a general argument, and then by trial and error finding the values that give tails of 0.025.

Example 4.8 continued...

The **mid-P limits** are given by

$$P_0 + P_1 + P_2 + P_3 + P_4 + 0.5P_5 = 0.975 \text{ or } 0.025$$

where P_i is the binomial probability (as in (2.9)) for i events with n = 20 and   = $_L$ or $_U$ This expression was set up on a personal computer for general   , and starting with the knowledge that the confidence interval would be slightly narrower than the limits of 0.0865 and 0.4908 found earlier the exact 95% mid-P confidence limits were found as 0.098 and 0.470. Method 2 gives the best approximation to these limits but, as noted earlier, the lower confidence limit is less well approximated by the normal approximation, because n $_L$ is only about 2.