Corella, 1995, 19(1): 25-34

SHEBA: COMPUTER PROGRAMS FOR SEXING BIRDS ON MEASUREMENTS USING UNIVARIATE DATA

K. G. ROGERS

340, Ninks Road, St Andrews, Victoria 3761

Received 26 March, 1994

This paper presents a suite of user-friendly computer programs designed for univariate analysis of measurements of sexually size dimorphic birds which have not been sexed by other means. Two of the programs estimate the measurement parameters for each sex (number, mean, standard deviation) and two use these estimates to produce a sexing criterion or rule by which a sex can be assigned to individual birds. Utility programs for preparing bivariate histograms are also presented. Although designed for the problem of sexing birds on measurements, the programs can be used for other univariate problems involving the separation of mixed normal distributions. The programs are only available for IBM machines (or clones) operating under DOS and are available from the Australian Bird Study Association, PO Box A313, Sydney South, NSW 2000.

OVERVIEW

Pyke and Armstrong (1993) described a method for assigning sex to birds of a species in which the sexes differ in size, with overlap between the sexes, but not in plumage. Rogers and Rogers (1995) expressed serious concerns with their approach but recognized that no convenient method was available to those who hold large data sets on such birds but who can do little with them without becoming statisticians. This paper describes a suite of user friendly computer programs, collectively called SHEBA, which help to fill this gap.

Many people, particularly bird banders, will have samples (often large) of measurements of birds for species which cannot be sexed on plumage. If the sexes differ in size, as is often the case, the measurements can be used to give an indication of the sex of individual birds. The histogram of a measurement for such a species will usually be double-humped (Rogers and Rogers 1995). SHEBA provides the means for developing a sexing criterion or rule from such data.

The first task is to separate the sexes i.e. to estimate the average size and the variability of size about this average for birds of each sex. The second task is to develop, using these parameters, the required sexing criterion. It may not be possible to sex all birds; this depends on the requirements of the analyst and how certain he wants to be of not assigning the wrong sex. A trade-off has to be made between the conflicting desires of sexing as many birds as possible and getting the sex wrong in as few cases as possible.

This is a difficult statistical problem and this paper does not deal with the statistical detail. Rogers *et al.* (1986) and Rogers and Rogers (1995) describe the relevant considerations. The statistical detail will be dealt with in papers on the more difficult problem of sexing birds using more than one measurement (Rogers, unpubl.). The programs are presented in an interactive format requiring the user to set the programs going and respond to questions on the computer screen.

The programs were designed for the problem of sexing birds and this paper confines discussion to this topic. They could be used for the analysis of other data sets, however generated, subject to the assumptions used. Before using them for sexing, the analyst must determine if an observed size dimorphism (i.e. a double-humped histogram) is due to sex or some other reason. This information will be available from the literature for many species; for others, examination of museum skins or cloacae of birds in breeding condition will help.

The only preliminary work required before the programs can be used is the preparation of a histogram of each measurement to be considered. Guidelines for doing this are given in Appendix A. In Appendix B, a general explanation of how the programs work is given.

The next section of the paper describes the programs. The following section illustrates their use and outputs. Some general comments concerning their application are made in the final section of the paper.

THE SHEBA PROGRAMS

SHEBA contains two programs for separating the sexes, two for developing sexing criteria, and two utility programs. Two versions of each program are provided. One version prints the output; the other sends the output to a computer file. The latter version is provided in case some combinations of computer and printer do not like the printer version.

Full operating instructions are provided with the programs. The programs, and the operating instructions, guide the user on the few choices that the programs request.

Separating the Sexes

The two programs are HUMPS and HUMPS_UV. Both estimate for each sex:

- the number of birds;
- the average value of the measurement; and
- the standard deviation of the measurement.

In addition, HUMPS_UV estimates how well these parameters have been estimated. It also produces better parameter estimates. The outputs enable a more robust sexing criterion to be developed than do those of HUMPS. Accordingly, HUMPS_UV is the preferred program but it does need good starting values to work (see below) and these can be provided by HUMPS.

HUMPS_UV requires starting estimates of the number of birds in the smaller sex and their average measurement, e.g. wing length, head bill (total head) length. The starting estimates are then successively improved automatically until no further improvement is possible. The program may not work if the starting estimates are not reasonably close to the final ones.

HUMPS works by evaluating a number of possibilities defined by a range of starting estimates provided by the user. The results are summarised on screen giving the analyst the opportunity to refine the ranges. This revision of the ranges can be continued as many times as needed to reach an answer of the required accuracy. The final estimates produced by HUMPS have always (so far) worked as starting estimates for HUMPS_UV.

HUMPS replaces an earlier version of the program which has been around for some time.

Sexing Criteria

The programs, CRIT and CRIT_UV, estimate the sexing criterion. Not all birds will be assigned a sex, the number that will depending on how confident the user wants to be of not assigning the wrong sex (see Rogers and Rogers 1995). The sexing criterion is presented as:

- the required minimum probability of correct sexing;
- the upper limit of size for birds of the smaller sex;
- the lower limit of size for birds of the larger sex.

The limits are expressed in terms of the precision with which the measurement concerned can be measured (e.g. 1 mm for wing length, 0.1 mm for head-bill length). Birds falling between the limits (i.e. in the grey zone of Rogers and Rogers 1995) will be unsexed. The programs also give the percentages of birds which will be sexed correctly, wrongly, or unsexed.

CRIT_UV is the preferred program as it uses the information from HUMPS_UV on the accuracy with which the parameters are estimated and the correlation between them. In particular, it accommodates the effect of sample size on the sexing criterion.

CRIT gives an approximate criterion which may be all that is possible if the user is using summary data, as from a handbook for example. This version replaces an earlier version of the program which has been around for some time.

Utility Programs

These are programs which the analyst might find useful in preparing data for input to HUMPS and HUMPS_UV.

HIST_BV compiles a bivariate histogram of any two measurements from a data file, e.g. wing length against head bill length. In this case, the program gives the number of observations which fall in every combination of wing length and head-bill length interval. The output allows for quick visual examination of the data for outlying or possibly wrongly measured data points. If any such points are identified, the analyst may choose to exclude them from the analysis. The program also compiles univariate histograms which can be used as input to HUMPS and HUMPS_UV.

MINMAX finds the minimum and maximum of each variable of the data file used as input to HIST_BV. These facilitate specification of the histograms.

USE AND OUTPUTS OF PROGRAMS

Data on head-bill lengths of 115 adult-plumaged Eastern Yellow Robins *Eopsaltria australis*, which were banded at several sites in Victoria, are analysed here. Table 1 gives the histogram of

TABLE 1

Histogram of Eastern Yellow Robin head-bill lengths.

	Mid-point of Histogram Interval (mm)														
	34.0	34.5	35.0	35.5	36.0	36.5	37.0	37.5	38.0	38.5	39.0	39.5	40.0	40.5	41.0
Observed frequency	1	7	10	9	10	7	7	17	12	17	12	4	0	1	1

Note: Lower limit of the first histogram interval is 33.75 mm. Interval size is 0.5 mm.

these measurements; these are the only data required to use the programs. Examination of museum skins (Rogers *et al.* 1986) showed that males are usually larger than females for this species. The histogram is not perfectly regular being based on a fairly small sample but shows clearly the double humped shape characteristic of data from a well separated sex size dimorphic species with the peaks of the humps at approximately 35.5 mm and 38 mm. The peaks are sufficiently close to suggest some overlap in size between the sexes.

Sexing Programs

Operation. The histogram is input to either HUMPS or HUMPS_UV. This can be done in one of two ways, either by providing information through the keyboard as requested by the programs or by preparing an input file of the data (using any text editor or word processing program). File entry is better and, ultimately, easier for the user. The user is asked to make a choice concerning the variability in size of the two sexes.

- either that the absolute variation in size about the average is the same for both sexes; or
- that the relative variation in size about the average is the same for each sex (implying a larger absolute variation for the larger sex).

The second choice seems more plausible than the first and examination of many data sets suggests that it is more generally applicable. With smaller birds and/or measurements, the first choice may give a better answer because the difference between the choices is small and may not be detectable, particularly with small samples. Also, measurer variation is more likely to mask any underlying pattern in small measurements. It would be good practice to run the programs using both choices to see which is most appropriate for a particular data set.

Both programs produce printed output and an output file. The output files of both can be used as input to both CRIT and CRIT UV; the latter will only work with the file from HUMPS UV. The user has the option for keyboard input. Again, the programs offer choices. The crucial ones are:

- how sure the user wants to be that the wrong sex is not assigned (the minimum probability of correct sexing); and
- how precisely the measurement can be recorded in the field. It would be silly to report a sexing criterion to three decimal places for a measurement that can only be recorded to the nearest millimetre (e.g. wing length).

CRIT_UV also needs information to calculate the accuracy of the final estimates. The user can use the default values provided or choose his own.

Outputs. Only the outputs of HUMPS_UV and CRIT_UV are described. These give the complete results. HUMPS and CRIT only provide part of them.

Figure 1 shows the outputs of HUMPS_UV using these data. Most of the entries in the table are self-explanatory. The table gives the observed and expected values in each histogram interval; clearly the fit is quite good. The area of overlap between the sexes, the grey zone of Rogers and Rogers (1995), lies between 35.25 mm and 38.25 mm. The difference between the means is large relative to the standard deviations, so a good sexing criterion can be expected. The standard errors are a measure of how well the parameters have been estimated; they are used in CRIT_UV in calculating the sexing criterion.

The programs require a measure that tells the analyst how good the answers are. Four are provided; see Appendix B for details. The final line gives the value of chi-squared, grouping histogram intervals so that all expected values are greater than or equal to 5. This value is comparable with chi-squared calculated by HUMPS.

The non-statistician can gain a sensible idea of how good the estimates are by simply comparing expected with observed values.

Figure 2 shows the output of CRIT_UV. The top part of the table repeats the input data estimated in HUMPS_UV.

If the measurement parameters (of wing, say) are known precisely, the probability that each sex will have a particular wing length is easily calculated. It is also easy, using these probabilities, to estimate the probability that a bird with this wing length is a male or a female. With a little bit more work, it is possible to find the values of wing beyond which the probabilities are greater than the required minimum probability of correct sexing. This is done in the program CRIT.

HUMPS_UV

Date: 04-07-1994

SPECIES: Eastern Yellow Robin
MEASURE: Head-Bill Length
AGE GROUP: Adult

EQUAL COEFFICIENTS OF VARIATION

ESTIMATES

	SMALLER SEX	LARGER SEX
NUMBER OF BIRDS	41	74
MEAN	35.416	38.148
S.D.	0.802	0.863

ASYMPTOTIC STANDARD ERRORS

A.S.E., SMALLER NUMBER	1.2060
A.S.E., SMALLER MEAN	0.0368
CORRELATION BETWEEN ESTIMATES	0.9218

OBSERVED AND EXPECTED VALUES

Interval Mid-Point	Observed Frequency	Estimated Frequency		
		All	Smaller	Larger
< next I'val	0	0.8	0.8	0.0
34.000	1	2.2	2.2	0.0
34.500	7	5.3	5.3	0.0
35.000	10	8.8	8.8	0.0
35.500	9	10.2	10.0	0.2
36.000	10	8.6	7.8	0.8
36.500	7	7.0	4.1	2.9
37.000	7	8.7	1.5	7.1
37.500	17	13.2	0.4	12.8
38.000	12	16.7	0.1	16.6
38.500	17	15.6	0.0	15.6
39.000	12	10.5	0.0	10.5
39.500	4	5.1	0.0	5.1
40.000	0	1.8	0.0	1.8
40.500	1	0.5	0.0	0.5
41.000	1	0.1	0.0	0.1
> last I'val	0	0.0	0.0	0.0

MAXIMUM LOG-LIKELIHOOD	-284.4616		
APPROXIMATE CHI-SQUARED	13.995	DEGREES OF FREEDOM	14
CHI-SQUARED (Observed>0)	15.742	DEGREES OF FREEDOM	14
CHI-SQUARED (Min expected=5)	3.888	DEGREES OF FREEDOM	10

Fig. 1. *HUMPS_UV output.*

CRIT_UV Date: 04-07-1994

SPECIES Eastern Yellow Robin
MEASURE Head-Bill Length
AGE Adult

INPUT DATA

	SMALLER	LARGER
NUMBER	41	74
MEAN	35.416	38.148
S.D.	0.802	0.863

A.S.E. SMALLER NUMBER	1.2060
A.S.E. SMALLER MEAN	0.0368
A. CORRELATION	0.9218

MONTE CARLO ESTIMATES

	COUNT	MEAN	S.D.	MIN	MAX
SMALLER SEX NUMBER	1000	40.9510	1.1463	37.000	44.000
SMALLER SEX MEAN	1000	35.4146	0.0372	35.277	35.519
SMALLER SEX NUMBER - A.S.E.	1000	1.2041	0.0789	0.920	1.441
SMALLER SEX MEAN - A.S.E.	1000	0.0368	0.0024	0.028	0.045
ASYMPTOTIC CORRELATION	1000	0.9211	0.0139	0.815	0.958

CONSEQUENTIAL ESTIMATES

SMALLER SEX S.D.	1000	0.8024	0.0132	0.761	0.853
LARGER SEX MEAN	1000	38.1467	0.0235	38.065	38.217
LARGER SEX S.D.	1000	0.8643	0.0138	0.822	0.917
UPPER LIMIT FOR SMALLER SEX	1000	35.9859	0.0401	35.836	36.104
LOWER LIMIT FOR LARGER SEX	1000	37.4812	0.0423	37.295	37.641
% RIGHT	1000	77.0462	1.7273	69.491	83.316

SEXING CRITERIA

SEX PROPORTIONS NOT USED IN CALCULATING CRITERION

MINIMUM CONFIDENCE LEVEL OF CORRECT SEXING 95.00 %

ADJUSTMENTS	UL SMALLER	LL LARGER	SEX	RIGHT	D.K.	WRONG
None	35.98586	37.48116	BOTH	77.06	22.38	0.56
			SMALLER	76.18	23.32	0.50
			LARGER	77.94	21.44	0.62
Sampling	35.91983	37.55077	BOTH	74.51	25.04	0.44
			SMALLER	73.55	26.06	0.39
			LARGER	75.47	24.03	0.50
Measurement	35.90000	37.50000	BOTH	75.01	24.52	0.47
			SMALLER	72.74	26.79	0.47
			LARGER	77.28	22.25	0.47
Both	35.90000	37.60000	BOTH	73.19	26.41	0.39
			SMALLER	72.74	26.94	0.32
			LARGER	73.65	25.89	0.47

Figure 2. *CRIT UV output.*

If, however, the measurement parameters are not known precisely, calculating the probability that a bird with a given wing length is a male is a difficult analytical task. It is tackled in CRIT UV by a Monte Carlo calculation, the results of which are given in the middle part of the output (Fig. 2).

Using the HUMPS_UV outputs, a large number of random samples of the parameter estimates are made. The simple CRIT method is applied to each sample and a number of estimates of the criterion limits obtained. Knowing the variability of the criterion limits, the sexing criterion can be set to ensure that sex is not assigned at less than the required confidence level. This correction to the expected criterion limits is the sampling adjustment. Finally, the measurement adjustment is made. This rounds the upper limit of size for the smaller sex down, and the lower limit of size for the larger sex up, to the nearest unit of measurement precision. Again, this is necessary to ensure that no birds are sexed at less than the required confidence level.

The Monte Carlo estimates (split into two parts, those sampled and those calculated using the sampled values) show that the input data have been well reproduced and suggest that the criteria presented in the bottom part of the table are likely to be robust.

How many samples should be taken in the Monte Carlo calculation cannot be generalized. The number will depend on sample size and the separation between the sexes. A good way to proceed is to take a small sample first, say 100, and see how well the input data are reproduced. Certainly, a fit to at least three significant figures is required. If this is not achieved, a larger sample is required. As a general rule, a sample of four times the size is needed to double the precision of the estimates.

The sexing criteria part of the output (Fig. 2) shows how well the sexing criteria perform. The top of this part of the table reports that sex proportions (those of the sample) were not used in calculating the criteria. This is appropriate if the criteria are to be used for sexing birds from outside the sample when the expected sex ratio is 50 per cent. If the sex ratio is known to be different from 50 per cent, as might be the case if the criterion were to be applied only to the sample birds for example, this prior knowledge should be built into the criterion. This option is available in the program.

```
MINMAX                                                  DATE 04-07-1994

Minima and Maxima of Variables on File - WPHE4VAR.PRN

  73.00  31.00  64.00  12.40
  93.00  37.50  88.00  26.20
```

Figure 3. *MINIMAX output.*

Sexing criteria are given if no adjustments are made and if the sampling and measurement adjustments are applied separately and together. Overall some 4 per cent fewer birds will be correctly sexable if both adjustments are made than if neither is. This result cannot be generalized as less well behaved data can lead to far bigger reductions in the percentage of birds which can be sexed with confidence. Note that, for these data, the sampling adjustment has no effect on the upper limit of size for the smaller sex. This will not always be the case.

Utility Programs

Data on adult White-plumed Honeyeater *Lichenostomus penicillatus,* caught in Victoria, are used to illustrate these programs.

Figure 3 gives the output of MINMAX. It simply gives the minimum and maximum of each variable on the input file.

Figure 4 gives an example of the output of HIST_BV, plotting Wing length against Head-Bill length. Similar plots could be obtained for any other combination of variables. The output throws up a number of suspect data points e.g. the one with a Head-Bill of 33 mm and a Wing of 89 mm. This would appear to be a female on Head-Bill and a male on Wing; one or both of these measurements would appear to be in error. The output (Fig. 4) also suggests that a bivariate sexing criterion is likely to be considerably more effective than a univariate one.

COMMENTS

A danger with producing easily used programs is that people will use them indiscriminately. If rubbishy data is put in, perhaps containing obviously suspect data points or poorly chosen histogram intervals, the programs might produce answers which look good but which are wrong. Used sensibly, the programs should help the analyst to gain an understanding of his data.

Histograms are not always as obviously double humped as that for the Eastern Yellow Robin data used here. if there sis considerable overlap

BIVARIATE HISTOGRAM for White-plumed Honeyeater DATE 04-07-1994
INPUT FILE wphe4var.prn

	Wing	by	Head-Bill
Lower Limit	72.5		31.75
Increment	1		.5

Observations 548
Corr. Coeff. .6240796

												Wing	
						1			1			2	93.00
					1			1		1		3	92.00
								4	3	2		9	91.00
				2	2		3	6	6	1		20	90.00
		1		1		6	7	11	12	2	2	42	89.00
			1		1	5	8	15	11	6	1	48	88.00
					3	8	12	3	9	3	2	40	87.00
			1	1	7	9	12	8	5	1		44	86.00
				5	2	6	16	3	9	3		44	85.00
			3	6	4	12	4	1	2	1		33	84.00
	1	2	2	11	8	6	5	2	2			39	83.00
	1		7	10	8	10	9	4	3	3		55	82.00
1	1	2	7	9	9	8	1	2	4	1		45	81.00
	2	4	7	7	8	7	4		1			40	80.00
		3	8	8	6	5						30	79.00
	1	2	5	1	3	1	1					14	78.00
		4	3	2	4	2						15	77.00
			3	3	2	1						9	76.00
	1	1	4	1	2	1	1					11	75.00
		1	1		1	1						4	74.00
							1					1	73.00
Head-Bill													
1	7	20	52	67	71	89	84	60	68	24	5		

32.00 33.00 34.00 35.00 36.00 37.00
32.50 33.50 34.50 35.50 36.50 37.50

Number of Histogram Intervals Head-Bill 12 Wing 21

Figure 4. *HIST BV output.*

between the sexes, the histogram might even be single humped. If it does not follow the clear bell shaped curve of the normal distribution (it might, for example, have a flat top or a bulge to one side), separation of the sexes might still be possible. Rogers and Rogers (1995) give an illustration of this.

Successful separation of the sexes does not necessarily mean that a useful sexing criterion can be obtained at a comfortable minimum level of confidence of correct sexing. This can happen if the variability of the measurements is large relative to the difference between the means; in this case, the grey zone will be large.

Use of the SHEBA programs does not prove that a species is sex size dimorphic; it does show that the sample can be considered as containing two groups of birds of different size. Independent evidence is needed to attribute that size difference to sex. Note that a sample may contain more than two humps; this might arise if, for example, it contains different subspecies or ages. In this case, some birds may have to be removed from the data before using SHEBA.

First time users of SHEBA might find it helpful to practice with the data used here (which is provided with the programs), first to repeat the results presented and, secondly, to gain experience with all the programs and choices available. Users who have difficulty with the programs or the documentation should contact me directly.

ACKNOWLEDGMENTS

Annie Rogers is to be thanked for telling me when she could understand this paper — it took several drafts. I am grateful too to Danny Rogers, Mark Barter and Andrew Dunn for comments on drafts of this paper and for testing the computer programs. My thanks too to the referees, Belinda Dettmann and Graham Fry, whose helpful comments were much appreciated.

APPENDIX A
HISTOGRAM DEFINITION

Defining a Histogram

A histogram is defined by the lower limit of the smallest (first) interval (or bin) and the size of the interval. Their selection should be considered together.

The lower limit of the first interval should obviously be less than the smallest observation. It should be specified to one more decimal place than the precision to which the measurement is recorded and, ideally, picked so that observations fall in the mid-point of the intervals.

For example, if the smallest head-bill length is 37.4 mm, set the lower limit of the first interval to 37.35 mm if an interval size of 0.1 mm is to be used; set the lower limit to 37.25 mm if an interval size of 0.5 mm is to be used.

The smaller the interval size, the number of intervals will be larger and the number of observations in the intervals will be smaller. It is not a good idea to have too many intervals with small (e.g. less than 10) numbers of observations.

More intervals may be supported by large samples and, to some extent, the range (difference between the largest and the smallest observations) of the sample. Measurement precision also comes into it. It is all very well measuring head-bill lengths to 0.1 mm but, if a bird were to be measured several times, would the same measurement always be recorded by different measurers or, for that matter, by the same measurer? It might be better in such a case to use intervals of 0.5 mm which will largely overcome the problem.

Ultimately, the analyst will have to make a not very hard subjective choice. Does the resulting histogram look sensible? If there are too few intervals, much of the data will be concentrated in a small number of them and the histogram will not show how most of the observations vary. If there are too many, the histogram will look like a saw tooth and may not show any pattern in the data. Generally, and this is a guide and not a rule, between 10 and 25 intervals will be found most useful.

On a practical point, it is a good idea to use the smallest interval size the data will support. Should examination show that the intervals are too small, it is a simple matter to 'lump' adjacent intervals to give a histogram with larger interval sizes. This saves going through the data for a second time.

Examples of Histogram Definition

The table below shows some fairly typical histogram definitions which can be used over measurements of different sizes and precision. Wings are measured to the nearest 1 mm, head-bills to the nearest 0.1 mm, weights (these are for large birds) to the nearest 5 gm.

Measure	Minimum	1st Maximum	Lower limit size Interval interval	Number of size	intervals
Wing	63	80	62.5	1	18
Head-Bill	36.0	42.3	35.75	0.5	12
Weight	325	395	322.5	5	15

APPENDIX B
MORE INFORMATION ON THE PROGRAMS

HUMPS

Apart from inputs and outputs, this is the same as the program made available in 1988 and used, at least, by Rogers *et al.* (1990), Barter (1985, 1986, 1989, 1990), Fry (1990) and for which a prototype was used by Rogers *et al.* (1986).

Given the number of birds in one sex, the number in the other is readily calculated. Given also an estimate of the mean of one sex, the mean of the other is readily calculated. Given the above, it is possible to calculate the standard deviation for each sex if a simplifying assumption is made concerning how the standard deviations of the sexes differ (this assumption is discussed in the body of the paper).

The user specifies a range of estimates of the number in, and mean of, one sex. Each range is split into four equal parts, giving five estimates to be evaluated for each parameter. HUMPS works by calculating the remaining parameters for each combination of estimates and the number of observations expected in each histogram interval. These are compared with actual observations using chi-squared. The program produces (on screen) a map of chi-squared over the ranges and asks the user if he (or she) wants to try different ranges. This process allows the ranges to be successively redefined as the user homes in on the values which give the minimum chi-squared.

HUMPS_UV

This program has the same theoretical basis as HUMPS but differs from it in two important respects. First, it searches for the parameter estimates which maximize the logarithm of the likelihood function (see, for example, Macdonald and Pitcher 1979). Maximum likelihood methods are preferred to chi-squared as, although both methods produce substantially the same estimates with large samples (Macdonald and Pitcher 1979), maximum likelihood methods also estimate their accuracy (asymptotic standard errors) and the relationship between them (asymptotic correlation). These parameters are used in calculating sexing criteria by CRIT_UV.

Secondly, HUMPS_UV is an optimizing routine; given a starting point (initial estimates of the parameters), it uses a systematic grid search to find the parameter values which give the maximum of the likelihood function. The grid search works by holding one of the initial parameter estimates steady and varying the other in discrete step sizes until no improvement in maximum likelihood is possible. This revised estimate is then held steady whilst the same process is repeated for the other estimate. That is then held steady and the process repeated for the first estimate. This to-ing and fro-ing is continued until no further improvement in maximum likelihood is possible.

HUMPS_UV has always worked when HUMPS values are used as the initial estimates. It could, however, fail to find the best answer if the data are badly behaved (for example, a saw tooth histogram with lots of possible local peaks and troughs). It is a good idea with optimising programs to use a number of different sets of initial estimates to ensure that a local optimum is not found.

The program requires three items of input data that HUMPS does not. The two step sizes used for the grid search have to be specified. The program can take a long time to run if small step sizes are specified initially. Things can be speeded up by starting off with large steps to find an approximate optimum and then running the program again with the approximate optimum estimates as initial values and using smaller step sizes. This is particularly necessary with large samples and many histogram intervals.

The final item relates to estimating the standard errors of the estimates and the correlation between them. This is done by examining how the likelihood function changes in a small range around the maximum; a single number defines 'small' as a proportion of the final estimates. The user may specify a required value or use the program default.

The standard errors calculated are asymptotic standard errors which will more closely approximate the true values to the extent that the likelihood surface is quadratic in the area of the maximum likelihood; this is generally considered to be a good assumption. Estimating the standard errors requires the second derivatives of the likelihood function with respect to the parameters being estimated (the Hessian matrix). Since it would be difficult to find these analytically, they are estimated by calculating the first derivatives by finite differences and using the Gaussian approximation to the Hessian to give the required second derivatives. A useful introductory text on non-linear estimation methods is Sadler (1975).

The last four lines of the HUMPS_UV output (Fig. 1) give information on how good the final estimates are. They are:

- MAXIMUM LOG-LIKELIHOOD. This is the function maximized by the program. It can be converted directly to an approximate estimate of chi-squared (Macdonald and Pitcher 1979), the commonly used statistic for measuring the correspondence between observed and expected values;
- APPROXIMATE CHI-SQUARED. Calculated from the maximum log-likelihood;
- CHI-SQUARED (Observed > 0). The expected number in each histogram interval is calculated using the final parameter estimates. Chi-squared is calculated over all histogram intervals containing at least one observation. The value should approximate the value in the line above, particularly for large samples;
- CHI-SQUARED (Min. expected = 5). This is the more usual chi-squared measure in which histogram intervals are lumped so that all cells used in the calculation have an expected value of 5 or greater. This is the chi-squared calculated in HUMPS.

CRIT

This program calculates a sexing criterion using the outputs of HUMPS and HUMPS_UV. The criterion is expressed as the upper limit of size for birds of the smaller sex (ULS) and the lower limit of size for birds of the larger sex (LLL) subject to a sex being assigned to a minimum probability of correct sexing specified by the analyst. It produces the same results as the program of the same name made available in 1988 and used by Rogers *et al.* (1986), but inputs and outputs have been changed. The coding has also been changed in one place for faster operation but answers are unaffected.

Using the estimated means and standard deviations, the probability that a male will have a particular wing length (say) is readily calculated from normal distribution theory. A similar probability can be calculated for a female with same wing length. It is not hard to calculate from these probabilities how much more likely a bird of unknown sex with that wing length is to be a male than a female. This can be generalized over all wing lengths that can occur. These figures can be examined to give theoretical values for ULS and LLL, the sexing criterion at any required minimum probability of correct sexing (it cannot be less than 50%). Two situations apply: if we have no reason to suppose that the sex ratio in the population is not 50:50 or if we have reason to believe that the ratio is imbalanced. The latter situation might apply to the sample birds if its estimated sex ratio differs substantially from 50 per cent; in this case a slight modification of the criterion is indicated.

The analyst has to select the minimum probability of correct sexing required. Convention would suggest 95 per cent for sexing individual birds — implying a never worse than 1 in 20 chance of assigning the wrong sex to a single bird. At this level, more than 95 per cent of birds will be correctly sexed since the value only applies at the limit. A minimum confidence level of 50 per cent would be appropriate if the analyst was only interested in the sex ratio. At this level, no birds will be unsexed but many will be wrongly sexed in the grey area

The program makes one adjustment to the theoretical criterion limits. ULS and LLL can be estimated to a greater degree of accuracy than the measurement can be recorded in the field. Accordingly, the program rounds ULS down, and LLL up, to the precision with which the measurement concerned can be recorded, e.g. wing lengths to the nearest millimetre, head-bill lengths to the nearest tenth of a millimetre. This ensures that no birds are assigned a sex with less than the required minimum probability of correct sexing.

The program will crash if the separation of the sexes is insufficient to support a criterion at the specified minimum probability of correct sexing.

CRIT_UV (see below) uses the standard errors of the estimates and the correlation between them to estimate the sampling error of the estimated confidence limits which can be used to define a more robust criterion; if this is done, the interval between the criterion limits will be wider than that produced by this program. CRIT has been included to enable a criterion to be obtained when only the distributional parameters are available as is the case with much published data e.g. many handbooks. The analyst will have to judge whether or not a criterion based on such data is meaningful. It should be if samples are large and sexes are well separated;

in this case standard errors of criterion limits are likely to be small leading to a small sampling adjustment.

CRIT_UV

There are two differences between this program and CRIT. First it makes a sampling adjustment to the theoretical criterion limits produced in CRIT which reflects how well defined are the parameters estimated by HUMPS_UV. The size of this adjustment will depend on the number of birds in each sex in the sample, the separation between the sexes, and the extent to which the data are consistent with normal distribution assumptions.

The second difference arises because it is impracticable (certainly difficult if possible) to calculate the adjustment analytically but it can be found by a Monte Carlo calculation. In this, a large number of random samples of the number of birds in the smaller sex, their mean, and the correlation between the estimates is taken and the sexing criterion found for each sample. This enables the standard deviations of ULS and LLL to be found and the appropriate limit consistent with the required minimum probability of correct sexing to be found. The sampling adjustment is made before the measurement precision adjustment.

This program can take a long time to run depending on the speed of the computer used, the data, and the number of Monte Carlo samples to be drawn. If your computer is very slow, or you want a very large number of samples, you might plan the use of CRIT_UV around mealtimes or bedtimes.

REFERENCES

Barter, M. A. (1985). Sex determination of bill length of live adult Curlew sandpipers. *Stilt* 7: 8–17.

Barter, M. A. (1986). Great Knots partly undone. *Stilt* 9: 5–20.

Barter, M. A. (1989). Bar-tailed Godwit in Australia. Part 1: Races, breeding areas and migration routes. *Stilt* 14: 43–47.

Barter, M. A. (1990). Morphometrics of the Eastern Curlew *Numenius madagascariensis*. *Stilt* 16: 36–42.

Fry, G. (1990). Biometrics, moult and migration of Terek Sandpiper *Tringa terek* spending the non-breeding season in north-west Australia. *Stilt* 17: 24–28.

Macdonald, P. M. and Pitcher, T. J. (1979). Age groups from size-frequency data: a versatile and efficient method of analysing distribution mixtures. *J. Fish. Res. Board Can.* 36: 987–1001.

Pyke, G. H. and Armstrong, D. P. (1993). Estimating sexes of New Holland Honeyeaters from head-bill measurements. *Corella* 17: 43–46.

Rogers, D. I. and Rogers, K. G. (1995). Commentary: estimating sexes of honeyeaters from head-bill measurement. *Corella* 19: 12–17.

Rogers, K. G., Rogers, A., Rogers, D. I., Lane, B. A. and Male, E. B. (1986). 'Bander's Aid: A Guide to Ageing and Sexing Bush Birds.' (A. Rogers: St Andrews, Victoria.)

Rogers, K. G., Rogers, A. and Rogers, D. I. (1990). 'Bander's Aid: Supplement Number One'. RAOU Report No. 67. (RAOU: Melbourne.)

Sadler, D. R. (1975). 'Numerical Methods for Nonlinear Regression.' (University of Queensland Press: St Lucia.)

RECOVERY ROUND-UP

This section is prepared with the co-operation of the Secretary, Australian Bird and Bat Banding Schemes, Australian Nature Conservation Agency. The recoveries are only a selection of the thousands received each year; they are not a complete list and should not be analysed in full or part without prior consent of the banders concerned. Longevity and distance records refer to the ABBBS unless otherwise stated. The distance is the shortest distance in kilometres along the direct line joining the place of banding and recovery; the compass direction refers to the same direct line. (There is no implication regarding the distance flown or the route followed by the bird). Where available ABBBS age codes have been included in the banding data.

Recovery or longevity items may be submitted directly to me whereupon their merits for inclusion will be considered.

Hon. Editor.

The following abbreviations appear in this issue:
AWSG — Australasian Wader Study Group.
NSW WSG — New South Wales Wader Study Group.
VWSG — Victorian Wader Study Group.

King Penguin *Aptenodytes patagonicus*

Y1494*. Adult (1+) banded on Île de la Possession, Crozet Islands, Indian Ocean (46°25'S, 51°45'E) in Dec. 90. Recovered, released alive with band on Heard Island, Antarctica (53°08'S, 73°43'E) on 13 Apr. 92. 1 735 km ESE.

*French Banding Scheme band.

Fiordland Penguin *Eudyptes pachyrhynchus*

J12130*. Adult (1+) banded at Taumaka Island, Open Bay Islands, New Zealand (43°52'S, 168°53'E) on 17 Nov. 93. Recovered dead at Shelly Point, Seamander, Tas. (41°28'S, 148°16'E) on 20 June 94. 1 702 km W.

*New Zealand Banding Scheme band.

Wandering Albatross *Diomedea exulans*

(a) 140–25115. Banded by J. D. Gibson at sea off Bellambi, NSW (34°20'S, 151°00'E) on 22 Aug. 64. Recaptured, released alive with band, on Adams Island, Auckland Islands (50°55'S, 166°00'E) on 26 Jan. 94, over 29 years 5 months after banding. 2 205 km SSE.

(b) 140–25801. Adult (1+) banded by S. G. Lane at sea of Malabar, NSW (33°58'S, 151°16'E) on 25 June 66. Recaptured, released alive with band, on Adams Island, Aucklands Islands (50°55'S, 166°00'E) on 7 Feb. 94, over 27 years 7 months after banding. 2 229 km SSE.