

HETFIT
Maximum Likelihood Estimates of
Heterogeneous Fecundability and Sterility

Darryl J. Holman
James W. Wood

Department of Anthropology
Population Research Institute
The Pennsylvania State University
University Park, PA 16802

This research was supported by NICHD Grants R01-HD20989 and R01-HD26747. The software and documentation were prepared for release with support from a predoctoral fellowship from the Population Council. Support was also provided by the Population Research Institute, Pennsylvania State University, which receives core support from NICHD grant 1-HD28263-01.

INTRODUCTION

Hetfit is a computer program that estimates multi-state models of fecundability and sterility from information on intervals from marriage to first live birth. The program is capable of estimating separate piecewise models over user-specified age ranges. Observations within each age interval are left-truncated, right censored, and dropped, as appropriate.

Hetfit models two distinct sources of heterogeneity. First fecundability is modeled as a constant hazard for each couple within a single age interval. However, fecundability can be made to vary according to a lognormal distribution among fecund couples. In addition, a discrete sterile state is modeled as a second form of unmeasured heterogeneity among couples. Sterility is modeled as a prevalence at the start of each age interval and as an incidence within an age interval.

The original application of Hetfit was to estimate effective fecundability from first birth intervals (Wood et al. 1994). However, the program can be used with appropriate data to estimate apparent fecundability or fecundability from higher order birth intervals. For example, apparent fecundability can be estimated from women coming off of contraceptives for whom pregnancy assays are given.

The Hetfit program is available at no charge via anonymous ftp over the internet. Compiled versions for MS-DOS, Windows, and Unix are available, as well as the Pascal source code for the program.

This document describes technical aspects of creating data sets for use with Hetfit, specifying models in the parameter file, running Hetfit, and obtaining Hetfit over the internet.. Some details of the numerical methods are included as well. Technical details of the model can be found in Wood et al. (1994).

USING HETFIT

The Hetfit program reads two files and creates one or more files for each run. The first steps in using Hetfit are to prepare a *data file* and a *parameter file*. These files are described in detail below.

Files Used by Hetfit

The Data File

This file contains a series of observations that are to be used to estimate model parameters. The requirements for this data file are very simple; however, the decisions that go into making up the data file can be quite complicated. Some of these issues are discussed below; however, a through reading of Wood et al. (1994) is essential.

Each line of the file contains a single observation. This observation corresponds to exposure, and possibly an event for one woman (or couple). Three numbers appear on each line, and the numbers are separated by one or more spaces or tabs. The data file is read once for each invocation of Hetfit.

Here is an example of a file:

The meaning of each column is:

262.395	262.695	1
263.975	264.275	1
230.018	230.318	0
268.875	269.175	1
271.154	271.454	1
231.219	231.519	0
287.214	287.514	1
.	.	.
.	.	.
.	.	.

1. The subject's age in months at marriage. This number can be an integer or a real value.
2. The subject's age in months when she experienced a fertile conception or the subjects age in months when she was right censored. This number can be an integer or a real value.
3. A *censored* code. This code is 1 if the observation is right censored and the code is 0 if the observation is a conception.

Note that when creating a data file from birth histories, the period of gestation must be subtracted off from the date of birth. Wood et al. (1994) drew a random period of gestation from a normal distribution with a mean of 8.6 months and a standard deviation of 1.15 months. However, other algorithms, such as subtracting a constant 8.6 months from each birth record, will also work.

For all observations, the subjects age at marriage must be *less than* the subject's age at last observation. When conception takes place at the same age as marriage, a fraction of a month should be added to the age at conception. For example, 0.033 months (about 1 day) can be used as the minimum time for conception or censoring.

The Parameter File

The parameter file defines several items such as the input and output file names and the specifications for the model. Details of the parameter file are given below. The parameter file may be read several times while Hetfit is running.

The Output File

Results of model runs are written to an output file. The name of this file is specified in the parameter file. Note that Hetfit will overwrite any output file that already exists with the same name.

Martingale Residual Output Files

Martingale residuals can be produced by Hetfit (see Wood et al. [1994] for details on martingale residuals and their interpretation). A single name for this file is specified in the parameter file, and Hetfit appends a file suffix for each age range for which Hetfit is run. A graphics program is then used to plot the residuals.

Running Hetfit

Hetfit is run from the DOS or Unix command line by typing

```
hetfit paramfile
```

where `paramfile` is the name of the parameter file. Hetfit has two optional arguments that produce extra debugging output. When any second command line argument is added to Hetfit, like

```
hetfit paramfile anything
```

Hetfit will write information useful for debugging the program or input data file. Most of this output will be sent to the standard output (the terminal). Some additional information is sent to the output file as well. When any third command line argument is added, Hetfit will produce even more debugging output. The debugging output is rarely required except when developing or modifying the program.

If Hetfit is run with more than three parameters, like

```
hetfit debug debug garbage
```

a long error message will be produced that summarizes the command line format, and the format of parameter files.

Hetfit can be run without specifying a parameter file name. Then, Hetfit will prompt the user for a parameter file name. This is particularly useful for running Hetfit under Windows.

Creating Models

Options to Hetfit are specified in the parameter file. The format of this file must be followed closely in order

for Hetfit to run. This section describes the parameter file format using the following example parameter file.

Specifying Input and Output Files

Many lines of the parameter file begin with a keyword, which must be typed in upper case characters exactly

```

INPUT_FILE    sample.dat
OUTPUT_FILE   sample.out
START_AGE     18           {Start at age 18}
END_AGE       26           {End at age 26}
INTERVAL_WIDTH 4           {There are 3 4-years intervals}
IGNORE_BELOW  18           {Throw away sub-18 year olds}
NUMBER_OF_MODELS 2       {full and no_hetero}
This is the full model
MARTIN fullmod
0.12  0.0000  1.1    0    Lambda_C
{Fecundability}
0.05  0.000  0.4    0    Sterility {Prevalence}
0.0   0.01   3.0    0    Sigma   {SD of z}
0.0001 0.00  1.0    0    Lambda_S {Incidence
sterility}
In this model, distribution of fecundabilities is turned
off
MARTIN OFF
0.12  0.02   1.1    0    Lambda_C
{Fecundability}

```

as shown (including any underscore characters). Input and output files are specified on the first two lines of the parameter file. The keyword INPUT_FILE is followed by one or more spaces or tabs, and followed by a filename. In this example the input file named SAMPLE.DAT will be read. Drive and path specifications are permissible up to 127 characters. Note that file extensions (e.g. .dat or .out) are required.

Specifying Age Intervals

The next five lines of the parameter file specify

```

INPUT_FILE    sample.dat
OUTPUT_FILE

```

the age intervals and the number of models to run. The keyword START_AGE specifies the age that opens up the first age interval. The age is specified in years. Note that all text after the 18 is ignored and can be used for comments (the brackets used around comments are not required).

The END_AGE keyword specifies the age for the last age interval. The INTERVAL_WIDTH keyword specifies how many years each interval will span. For this example, three separate runs of each model will be run in turn. The first will use the age interval [18, 22), the second [22, 26), and the third from [26, 30).

The IGNORE_BELOW keyword specifies a marriage age below which observations will be dropped. This should be equal to (or less than) the START_AGE.

Creating a Model

The keyword NUMBER_OF_MODELS specifies how many different models will be specified in the rest of the parameter file. In this example, two models will be specified.

```

START_AGE     18           {Start at age 18}
END_AGE       26           {End at age 26}
INTERVAL_WIDTH 4           {There are 3 4-years
intervals}
IGNORE_BELOW  18           {Throw away sub-18 year

```

For each model specified, there are six additional lines in the parameter file. The first line is a comment or title line. This line is written to the output to aid in identifying what model the output refers to. The MARTIN keyword

NUMBER_OF_MODELS	2			{full and no_hetero}	
This is the full model					
MARTIN	fullmod				
0.12	0.0000	1.1	0	Lambda_C	
{Fecundability}					
0.05	0.000	0.4	0	Sterility	{Prevalence}
0.0	0.01	3.0	0	Sigma	{SD of z}
0.0001	0.00	1.0	0	Lambda_S	{Incidence}

specifies the base name to use in producing martingale residual files, and is discussed in more detail below. The next four lines specify names, limits, starting values, and restrictions for each of the four parameters of the model. Each line has 5 columns (and then optional comments). The meaning of each column is:

1. This column specifies the starting value for the hill-climber to use for free parameters. The value is merely a starting guess that is likely to be near the final estimate. However, for all parameters that are held constant (reduced models), this column specifies the value of the constant parameter.
 2. This column specifies the lower limit of the range over which the hill-climber will search for a maximum. This column has no meaning for fixed parameters (but must be there nonetheless). For all free parameters, the value in this column should be greater than or equal to zero, since negative hazards, probabilities, and standard deviations are meaningless.
 3. This column specifies the upper limit of the range over which the hill climber will search for a maximum. For fixed parameters, this column has no meaning. The value in this column should always be greater than the value in the previous column. For the first, third and fourth parameters (fecundability, standard deviation of fecundability, and incidence of sterility) there is no mathematical constraint on the upper limit. The second parameter (prevalence of sterility) is a probability and should be constrained one or less. Note that the parameter limits shown in this parameter file are quite broad. Narrower limits can be used if difficulties of convergence or difficulties in finding the global maximum are encountered.
 4. This column specifies a *hold* on parameters. When *hold* is zero, the parameter is free for estimation by Hetfit. When *hold* is one, the model is reduced, and the parameter is fixed to the value found in the first column.
1. This column specifies a name for the parameter. Note that the order of parameters is fixed by the program, but the user can define any name whatsoever for parameters. Using the notation of Wood et al. 1994, the first parameter is λ_c , the second is s , the third is σ , and the fourth is λ_s .

For the reduced model below, the third parameter, σ (named Sigma here), is held constant at zero. This reduces the model to a form lacking a distribution of heterogeneity in fecundability among couples. In a similar way, values of λ_c , s , and λ_s can be fixed. Any number of parameters can be held constant in this way.

Producing Martingale Residuals

The MARTIN statement has one of two forms, seen in the box below. The root name of the martingale

```

In this model, distribution of fecundabilities is turned
off
MARTIN OFF
0.12  0.02  1.1  0  Lambda_C
{Fecundability}
0.00  0.00  0.4  0  Sterility  {Prevalence}
0.0  0.01  3.0  1  Sigma  {SD of z}
```

residual files is specified in the first line. These files will have the age-interval appended to them. So, for the above parameter file, the files *fullmod.18*, *fullmod.22*, and *fullmod.26* will be produced with one residual corresponding to each observation that contributed to the age group.

The second form in the box below specifies that no martingale residual file will be produced. The keyword OFF must be in upper case letters.

Producing Likelihood Surfaces

Some data sets (particularly small ones or those with a series of implausible observations) produce irregular ridges or many local maxima in the likelihood surface. Thus, plotting the likelihood surface can be a useful diagnostic tool. The Hetfit parameter file can be set to produce values for a three dimensional likelihood surface. The structure of the parameter file is modified when points for a likelihood surface are generated.

```

MARTIN
fullmod
MARTIN
```

The first major difference in this parameter file is that the NUMBER_OF_MODELS to run is set to zero. This tells Hetfit that the rest of the file specifies information for generating a likelihood surface. The next line, NUMBER_POINTS, specifies the number of points along x and along y that will be generated. Thus, in this file there will be ten x and ten y values for a total of 100 points generated.

```

INPUT_FILE  sample.dat
OUTPUT_FILE sample.lik
START_AGE   18
END_AGE     18
INTERVAL_WIDTH 11
IGNORE_BELOW 18
NUMBER_OF_MODELS 0  {Generate likelihood surface
instead}
NUMBER_POINTS 10  {Specifies the grid in x and
y}
0.12  0.05  0.50  0  Lambda_C  {This will
be x}
0.05  0.025  0.25  0  Sterility  {This will
```

The next four lines take on new meanings, but depending on the value of the *hold* switch in column 4. The *hold* switch now means that the parameter will be fixed to a value (*hold*=1) or that the parameter will be one of the axes of the surface (*hold*=0). Only the first two parameters with *hold*=0 are taken for producing the surface.

When *hold*=0, the columns, in order, mean: (1) nothing, (2) the starting value along the axis, (3) the largest value along the axis, (4) the *hold* value, and (5) the parameter name. The rest of the line is ignored and can be used for comments. When *hold*=1, the columns, in order, mean: (1) the value to fix this parameter at, (2) nothing, (3) nothing, (4) the *hold* value, and (5) the parameter name.

In the above example, a likelihood will be generated at 100 points. The x axis will reflect changes in

Lambda_C as 0.05, 0.10, ..., 0.5. The y axis will reflect changes in Sterility as 0.025, 0.5, ..., 0.25. The z axis will be the log likelihood at the particular value of x and y , Sigma = 0, and Lambda_S = 0.0001.

The following SAS program reads and displays the likelihood surface (note that hardware-specific code for `goptions` has been truncated). The likelihood surface file will contain a few lines of initial text that should be removed when using the SAS program.

Interpreting Results of hetfit Runs

The following is an annotated partial output from the sample parameter file included with Hetfit (which is not the same as the sample parameter file above).

```
filename gsasfile pipe 'lpr -Plp2';
goptions reset=global /*reset all to
defaults*/
      display device=xbw /*disp. dev. use xbw or
xcolor*/
...;
run;
data one;
  infile 'surface15.out';
  input lambda s llik;
  llik = 1000 - llik;
proc g3d;
  plot lambda*s=llik/rotate=80 tilt=80;
```

```

Start_age = 216 months.
End_age = 264 months.
Width of each interval is 48 months.
Marriages: Ignore_below 216 months of age.
Input file for this run was: "sample.dat"

=====
Working on interval 18 to 22
Observations not left truncation and failed: 89
Observations not left truncation and censored: 25
Observations left truncation and failed: 0
Observations left truncation and censored: 0
Observations dropped or bad: 0
Running 4 different models.

=====
==== Working on Model 1 ====
No heterogeneity Model. Sigma is fixed at 0
Martingale residual file is OFF

After 5 iterations, ln(L) = -217.664
This is the log likelihood for this model.

Parameter Name      Class      Parameter Estimate      Std Err
Lambda_C            free      0.2446850741           0.02637
Sterility            free      0.0123659486           0.09162
Sigma                fixed     0.0000000000           -----
                    sigma is fixed at 0
Lambda_S            free      0.0000000001           0.04306
                    Estimate of  $\lambda_c$  and SE
                    Estimate of  $\sigma$  and SE
                    Estimate of  $\lambda_s$  and SE

Variance/covariance matrix:
0.00069545      0.00004547      -0.00011795
0.00004547      0.00839457      -0.00216050
-0.00011795     -0.00216050     0.00061468
.
.
.
.
.
Skip to model four in this age group
.
.
.
==== Working on Model 4 ====
Full 4 parameter model. All parameters are free
Martingale residual file is OFF

After 3 iterations, ln(L) = -214.106
Use this likelihood with the one above for a likelihood ratio test

Parameter Name      Class      Parameter Estimate      Std Err
Lambda_C            free      0.2726274770           0.06742
Sterility            free      0.0000000001           0.10013
Sigma                free      0.7700953140           0.44476
Lambda_S            free      0.0000000001           0.02511
All parameters are free now

Mean fecundability = 0.36659 Variance in fecundability = 0.1073352

Variance/covariance matrix:
0.00454585      -0.00102074      -0.01726098      0.00006225
-0.00102074     0.01002587      -0.01221898      -0.00182537
-0.01726098     -0.01221898     0.19781583      -0.00146255
0.00006225      -0.00182537     -0.00146255     0.00063043

=====
Working on interval 22 to 26
Now do all four models over again for the next age group.
Observations not left truncation and failed: 61
.
.
.
.
.

```

The output is largely self explanatory. Note that an estimate of the mean and variance in fecundability is produced when σ is greater than zero; when it is zero, the mean fecundability is equal to λ_s .

For each age interval, Hetfit classifies each observation as one of 5 types:

Observations not left-truncated and failed: these are observations for which both marriage and a fertile conception is observed in the current interval.

Observations not left-truncated and censored: these are observations in which marriage began in the current interval, but the observation was right censored in or at the end of the interval.

Observations left-truncated and failed: these are observations in which marriage began prior to this interval and a fertile conception is observed in the current interval.

Observations left truncation and censored: these are observations in which marriage began in a previous interval and the subject is right censored in or at the end of the current interval.

Observations dropped or bad: These are observations not included in the current interval because: (1) The current age interval begins at an age later than the last time the subject was observed. In other words, the subject was right censored or had a fertile conception in a previous interval. (2) the current age interval ends before the age at which the subject gets married. (3) The age at last observation (whether censored or a conception) begins at or before the age that opens the interval (i.e. $t_i = r_i$ from equation 2 in Wood et al. 1994:410). (4) The age at last observation (whether censored or a conception) begins at or before the age of marriage in the current interval (again, this implies $t_i = r_i$). As noted previously, a small fraction of a month should be added to observations in which conception took place immediately after marriage.

ERROR MESSAGES

Hetfit provides WARNINGS, which do not terminate the program, and ERRORS, which terminate the program. ERRORS are printed as the first line of a long message that describes the format of the parameter file.

WARNING: Something is wrong with input file, line: xxx. Skipped

Hetfit provides simple checks of the input file. This message is printed when an input file line fails these checks. The checks are (1) that the date of marriage is earlier than the date of conception or censoring (conception must occur after marriage), (2) the dates are negative, or (3) the censor index is not 0 or 1.

WARNING: The matrix is singular

This warning suggests that the variance-covariance matrix is singular. The standard errors are not reliable.

WARNING: Max iterations (200) reached in dimension: x

This warning indicates that Hetfit had difficulty converging in the dimension (parameter) specified by x . Estimates are unlikely to be reliable when this message is produced. This problem can usually be fixed by specifying narrower or more reasonable ranges of parameters over which to hunt for a maximum. Note that DOS versions have more difficulty with this than do Unix versions.

WARNING: Max iterations (30) reached over all dimensions

This warning indicates that Hetfit did not converge after 30 iterations of the hill climber. Parameters might be close to a good value. Try using new starting values, or more reasonable or narrower ranges for parameters. As a last result, recompile Hetfit with a larger number of iterations specified.

ERROR: End of parameter file *parameter_file* was found too soon.

This error suggests an improperly specified parameter file. Perhaps the NUMBER_OF_MODELS has more models specified than are included. Six additional lines are required for *each* model.

ERROR: Didn't find the input file *input_file*.

The input file specified in the parameter file was not found. Make sure the extension is properly listed, and the directory and drive are properly specified, if necessary.

ERROR: Error converting *xxxx* into an integer.

This error is printed when Hetfit expects to read an integer value from the parameter file, and a non-integer string of text was found. Make sure there is no decimal point or extraneous text in the string listed as *xxxx*.

ERROR: Unknown token found in paramfile: xxxxxx

This message is printed when Hetfit finds a bad keyword in the parameter file. Make sure there are no lines missing from the parameter file.

ERROR: Missing NUMBER_OF_MODELS token

ERROR: Missing MARTIN token

ERROR: Missing NUMBER_POINTS token

Hetfit could not find this token. Make sure the order and spelling are correct for the line with this keyword.

ERROR: Could not find parameter file *paramfile_name*

This error message is printed when Hetfit cannot find the parameter file named on the command line. Check the spelling, path, and drive of the filename you specified.

ERROR: Usage: Hetfit param_file_name [debug [debug2]]

This error message is printed when the command line is improperly specified. Note that there can be one, two, or three command line parameters. The first is the parameter file name. The other two turn on generation of debugging output.

TECHNICAL NOTES

Compiling Hetfit

Hetfit was originally written in Turbo Pascal v. 3 for MS DOS. However the current DOS version of Hetfit was compiled in Borland Pascal 7.0 in real mode and the Windows version was compiled by Borland Pascal 7.0 in Windows mode. The current version can also be compiled under in most Unix systems that have a standard c compiler. However, to do this, the Pascal program is first translated into c using a Pascal-to-c translator called *p2c*. This translator is available from a number of ftp sites. Because the p2c translator can be compiled using most reasonable c compilers on a large number of computer architectures and operating systems, Hetfit can be made to run on many systems. The Pascal source code for the DOS and Unix versions of Hetfit are identical.

Hardware Requirements

The DOS version of Hetfit requires an Intel 80386 CPU or higher with a math coprocessor. However, the program can be compiled to run on 8086- and 80286-based machines and systems without math coprocessors by changing some compiler switches and recompiling the program.

Hetfit dynamically allocates memory to store observations as it reads in the data file. Each observation requires four extended precision real numbers, one long integer, one byte and a far pointer. This corresponds to 41 bytes per observation.

Hetfit is a CPU intensive program, and solutions for large data sets can require an enormous amount of computer time. Models that estimate a distribution of unmeasured heterogeneity in fecundability are particularly slow to solve. The following table shows timings of Hetfit runs on different hardware configurations using the SAMPLE.DAT input file (200 observations) and SAMPLE.PF parameter file. Timings on the Intel system were under DOS 6.2. Timings on the Sun systems were on lightly loaded machines.

Notes on Numerical Methods

Likelihood Maximizer (Hill-Climber)

Hardware	Time (secs)
Intel 80486/DX4-100 (DOS)	235.5
Sun SPARC 2	123.3
Sun SPARC 10/51	44.3
Sun SPARC 1000	36.0

A number of numerical hill climbing routines (used to find the global maximum likelihood) were tried for Hetfit. We did not want to use a routine that required analytical solutions for the partial derivatives of the likelihood with respect to each parameter. A simplex hill-climber was used in early versions of the program, using the algorithm in Press et al. (1989:326). For

the full model, this method was not reliable at finding the global maximum. Instead, a multidimensional hill climber was constructed that works in one dimension at a time (what Nelson [1982] calls the direct search method). Each iteration involves separately finding the maximum in each dimension. With each iteration through all dimensions, the convergence criterion is made more stringent. Hetfit uses a one-dimensional Brent parabolic interpolation hill climber for this purpose (Press et al. 1989:318). The resulting multidimensional hill-climber used in Hetfit is slow, but appears to work well at finding the global maximum.

Estimates of Standard Errors

Standard errors are estimated as the inverse of the Fisher's information matrix using parameter estimates in place of the true values. The method estimates the second partial derivative of the likelihood with respect to each parameter using the first term of a Taylor's series expansion and the first partial derivative (see Nelson 1982:349). First partial derivatives are computed by perturbation. For each dimension, a suitable δx is iteratively found that changes the likelihood by no more than 0.01 (to minimize error due to numerical overflow) and no greater than 0.1 (to minimize error due to a large δ). The information matrix is inverted by Gauss-Jordan elimination using the algorithm of Wang (1986:chap 2).

Integration

Hetfit uses very simple routines for numerical integration. A fourth order open-end extended trapezoidal rule integrator is used following Press et al. (1989:123, eq. 4.1.8). Single integrals and the inner of double integrals use 30 points with an expected precision of order 30^{-4} or about 1×10^{-6} . The outer of double integrals uses 25 points with an expected precision of order 25^{-4} or about 3×10^{-6} .

HOW TO OBTAIN HETFIT ON INTERNET

Hetfit is available via anonymous ftp from the Population Research Institute at the Pennsylvania State University (<ftp.pop.psu.edu>). Enter the login *anonymous*, and your email address as a password.

Once you are logged in, change into the pop directory by typing "cd pop". Change into the Hetfit directory by typing "cd hetfit". Be sure to set your ftp software to receive binary files.

Two Hetfit archives are available. The first is a self-extracting DOS archive called HETFITAR.EXE. The second is a Unix compressed tar archive called HETFIT.TAR.Z. Both archives contain the same files, listed below. Note that the p2c (Pascal to c) translator used to compile the Unix version of Hetfit is available at this ftp site as well.

Files

The following files are included in the Hetfit archive.

hetfit.exe	The DOS executable program
hetfitw.exe	The Windows executable program
hetfit	Sun OS executable program
hetfit.pas	The Pascal source code for Hetfit (DOS, Windows and Unix)
sample.dat	A file of fictitious data
sample.pf	An example parameter file that estimates parameters from sample.dat
samplelik.pf	An example parameter file that creates points for a likelihood surface
Makefile	A Unix makefile for Hetfit
hetfit.rtf	An RTF version of this document (some formatting may be lost).
surface.sas	A sample SAS program to generate likelihood surfaces.
hfred.me	A document with preliminary instructions.

Extracting Files From the Archives

The way in which you unarchive the Hetfit files depends on what type of computer you are using.

DOS

If you uploaded the file HETFITAR.EXE to a DOS- or Windows-based computer, extract the files as follows:

1. Create a directory to use for the extract (e.g. MD C:\HETFIT).
2. Move the HETFITAR.EXE file into the directory (e.g. MOVE HETFITAR.EXE C:\HETFIT).
3. Change into the extract directory (e.g. CD C:\HETFIT).
4. Execute the archive by typing: HETFITAR. The files listed above will be extracted.
5. The Windows version of Hetfit HETFITW.EXE, can be run using the "File-->Run" menu, or you can set up a program item so you can execute HETFITW.EXE by clicking on an icon.
6. If you do not want the Sun executable version of Hetfit, delete the file: hetfit

Unix

If you uploaded the file hetfit.tar.Z to a Sun or other Unix computer, extract the files as follows:

1. Create a directory to use for the extract (e.g. mkdir ~/hetfit).
2. Move the hetfit.tar.Z file into the directory (e.g. mv hetfit.tar.Z ~/hetfit).
3. Change into the extract directory (e.g. cd ~/hetfit).
4. Uncompress the archive by typing: uncompress hetfit.tar.Z
5. Untar the archive by typing tar -xvf hetfit.tar
6. If you do not want the DOS and Windows versions of Hetfit, delete the files: hetfit.exe and hetfitw.exe.
7. The executable file, hetfit, will run under Sun OS.

PROBLEMS AND DISCLAIMER

Hetfit is a complicated program that undertakes the imprecise art of numerically finding maximum likelihood estimates. We have tested Hetfit and believe it to be reasonably free of errors. However experience with software dictates that errors remain. If you have difficulties or find bugs in Hetfit, please let us know. We will try to eliminate errors or provide guidance on getting the program to work for your application. However, we cannot provide a warranty that Hetfit will work for any particular application.

If you have difficulties or questions, contact Darryl Holman by email: holman@pop.psu.edu

REFERENCES

- Wood JW, Holman DJ, Yasin A, Peterson RJ, Weinstein M, Chang M-c (1994) A multistate model of fecundability and sterility. *Demography* 31(3):403-426.
- Nelson W (1982) *Applied Life Data Analysis*. New York: Wiley.
- Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1989) *Numerical Recipes in Pascal: The Art of Scientific Programming*. Cambridge, UK: Cambridge University Press.
- Wang CK (1986) *Structural Analysis on Microcomputers*. New York: MacMillan

