NAME

HIST − Compute the sample probability density and inverse cumulative distribution of data.

SYNOPSIS

CALL HIST(X,N,LPDF, NC, PDF,CDFI,RC)

X(N)	is the REAL*8 vector of data values to be histogrammed
N	is the INTEGER*4 number of data values in X
LPDF	is the INTEGER*4 leading dimension of PDF and CDFI
NC	is the INTEGER*4 number of histogram cells used; see below
PDF(LPDF,2)	is the REAL*8 tabular probability density returned
CDFI(LPDF,2)	is the REAL*8 tabular inverse cumulative distribution function returned
RC	is the INTEGER*4 return code; see below

DESCRIPTION

If NC=0 on input, empty cells are not permitted. If NC ≤ 0 on input, NC is changed on return to the number of histogram cells actually used, as determined by Sturges' rule, NC = 1+ceil(lg N). Using NC and the maximum and minimum data values in X, the routine finds the cell width and boundary coordinates. Then it accumulates the cell frequencies in PDF. If a cell is empty and that has been disallowed, NC is reduced, new cell boundaries are found, and the cell frequencies are accumulated again. Then the cell frequencies are integrated (summed) to obtain CDFI, and both PDF and CDFI are normalized to make CDFI(NC+1,1)=1. The resulting PDF can be displayed using PRNHST, and the resulting CDFI can be used with INTERP to transform uniformly-distributed random variates into variates having the probability distribution of the data in X.

SEE ALSO

PRNHST, which prints a histogram
INTERP, which interpolates in a table

DIAGNOSTICS

These are the values of RC that can be returned.

0 all went well
1 N < 2 or LPDF < 3 or LPDF < NC+1
2 the cell width is zero

LINKAGE

gfortran source.f −L${HOME}/lib −lmisc

AUTHOR

Michael Kupferschmid

EXAMPLE

```
REAL*8 X(1000),PDF(12,2),CDFI(12,2)
INTEGER*4 RC
CALL DR250(1000,X)
NC=-1
CALL HIST(X,1000,12, NC, PDF,CDFI,RC)
PRINT *,'RC=',RC,' NC=',NC
MARGIN=50
CALL PRNHST(6,'DR250 output',12,1000,PDF,12,NC,MARGIN,RC)
STOP
END
```

In this example N=1000 so lg(N)=9.97 and Sturges' rule sets NC=1+ceil(9.97)=11; thus LPDF=12 is an adequate leading dimension for PDF and CDFI. Running the program produces the following output:

```
unix[1] a.out
 RC=           0  NC=          11
 DR250 output
 1.16E-03 ..................................
 9.19E-02 ..................................
 1.83E-01 ..............................
 2.73E-01 ..............................
 3.64E-01 ..................................
 4.55E-01 ..............................
 5.45E-01 ..............................
 6.36E-01 ..................................
 7.27E-01 ......................................
 8.18E-01 ..................................
 9.08E-01 ..................................
 9.99E-01
unix[2]
```

From this histogram it appears that the distribution of random numbers returned by DR250 is approximately uniform on [0,1].