NAG Library Routine Document

G07ABF

Note: before using this routine, please read the Users' Note for your implementation to check the interpretation of ***bold italicised*** terms and other implementation-dependent details.

1 Purpose

G07ABF computes a confidence interval for the mean parameter of the Poisson distribution.

2 Specification

```
SUBROUTINE G07ABF (N, XMEAN, CLEVEL, TL, TU, IFAIL)

INTEGER              N, IFAIL
REAL (KIND=nag_wp)   XMEAN, CLEVEL, TL, TU
```

3 Description

Given a random sample of size n, denoted by $x_1, x_2, \ldots, x_n$, from a Poisson distribution with probability function

$$p(x) = e^{-\theta}\frac{\theta^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

the point estimate, $\hat{\theta}$, for θ is the sample mean, $\bar{x}$.

Given n and $\bar{x}$ this routine computes a $100(1-\alpha)\%$ confidence interval for the parameter θ, denoted by $[\theta_l, \theta_u]$, where α is in the interval $(0, 1)$.

The lower and upper confidence limits are estimated by the solutions to the equations

$$e^{-n\theta_l}\sum_{x=T}^{\infty}\frac{(n\theta_l)^x}{x!} = \frac{\alpha}{2},$$

$$e^{-n\theta_u}\sum_{x=0}^{T}\frac{(n\theta_u)^x}{x!} = \frac{\alpha}{2},$$

where $T = \sum_{i=1}^{n} x_i = n\hat{\theta}$.

The relationship between the Poisson distribution and the χ^2-distribution (see page 112 of Hastings and Peacock (1975)) is used to derive the equations

$$\theta_l = \frac{1}{2n}\chi^2_{2T,\alpha/2},$$

$$\theta_u = \frac{1}{2n}\chi^2_{2T+2,1-\alpha/2},$$

where $\chi^2_{\nu,p}$ is the deviate associated with the lower tail probability p of the χ^2-distribution with ν degrees of freedom.

In turn the relationship between the χ^2-distribution and the gamma distribution (see page 70 of Hastings and Peacock (1975)) yields the following equivalent equations;

$$\theta_l = \frac{1}{2n}\gamma_{T,2;\alpha/2},$$

$$\theta_u = \frac{1}{2n}\gamma_{T+1,2;1-\alpha/2},$$

where $\gamma_{\alpha,\beta;\delta}$ is the deviate associated with the lower tail probability, δ, of the gamma distribution with shape parameter α and scale parameter β. These deviates are computed using G01FFF.

4 References

Hastings N A J and Peacock J B (1975) *Statistical Distributions* Butterworth

Snedecor G W and Cochran W G (1967) *Statistical Methods* Iowa State University Press

5 Parameters

1: N – INTEGER *Input*

On entry: n, the sample size.

Constraint: $N \geq 1$.

2: XMEAN – REAL (KIND=nag_wp) *Input*

On entry: the sample mean, $\bar{x}$.

Constraint: $XMEAN \geq 0.0$.

3: CLEVEL – REAL (KIND=nag_wp) *Input*

On entry: the confidence level, $(1-\alpha)$, for two-sided interval estimate. For example CLEVEL $= 0.95$ gives a 95% confidence interval.

Constraint: $0.0 < CLEVEL < 1.0$.

4: TL – REAL (KIND=nag_wp) *Output*

On exit: the lower limit, θ_l, of the confidence interval.

5: TU – REAL (KIND=nag_wp) *Output*

On exit: the upper limit, θ_u, of the confidence interval.

6: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction for details.

For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, if you are not familiar with this parameter, the recommended value is 0. **When the value -1 or 1 is used it is essential to test the value of IFAIL on exit.**

On exit: IFAIL $= 0$ unless the routine detects an error or a warning has been flagged (see Section 6).

6 Error Indicators and Warnings

If on entry IFAIL $= 0$ or -1, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

IFAIL $= 1$

On entry, N < 1,
or XMEAN < 0.0,
or CLEVEL ≤ 0.0,
or CLEVEL ≥ 1.0.

IFAIL $= 2$

When using the relationship with the gamma distribution to calculate one of the confidence limits, the series to calculate the gamma probabilities has failed to converge. Both TL and TU are set to zero. This is a very unlikely error exit and if it occurs please contact NAG.

7 Accuracy

For most cases the results should have a relative accuracy of $\max(0.5\mathrm{E}{-12}, 50.0 \times \epsilon)$ where ϵ is the ***machine precision*** (see X02AJF). Thus on machines with sufficiently high precision the results should be accurate to 12 significant digits. Some accuracy may be lost when $\alpha/2$ or $1 - \alpha/2$ is very close to 0.0, which will occur if CLEVEL is very close to 1.0. This should not affect the usual confidence intervals used.

8 Further Comments

None.

9 Example

The following example reads in data showing the number of noxious weed seeds and the frequency with which that number occurred in 98 sub-samples of meadow grass. The data is taken from page 224 of Snedecor and Cochran (1967). The sample mean is computed as the point estimate of the Poisson parameter θ. G07ABF is then called to compute both a 95% and a 99% confidence interval for the parameter θ.

9.1 Program Text

```
    Program g07abfe

!     G07ABF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: g07abf, nag_wp
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter                :: nin = 5, nout = 6
!     .. Local Scalars ..
      Real (Kind=nag_wp)                :: clevel, sum, tl, tu, xmean
      Integer                           :: ifail, ifreq, n, num
!     .. Intrinsic Procedures ..
      Intrinsic                         :: real
!     .. Executable Statements ..
      Write (nout,*) 'G07ABF Example Program Results'
      Write (nout,*)
```

```
!     Skip heading in data file
      Read (nin,*)

!     Read in counts and frequencies
      sum = 0.0E0_nag_wp
      n = 0
d_lp: Do
        Read (nin,*,Iostat=ifail) num, ifreq
        If (ifail/=0) Then
          Exit d_lp
        End If

!       Calculate sum
        sum = sum + real(num,kind=nag_wp)*real(ifreq,kind=nag_wp)
        n = n + ifreq
      End Do d_lp

!     Estimate poission parameter
      xmean = sum/real(n,kind=nag_wp)
      Write (nout,99999) 'The point estimate of the Poisson parameter =', &
        xmean
      Write (nout,*)

!     Calculate 95% confidence interval
      clevel = 0.95E0_nag_wp
      ifail = 0
      Call g07abf(n,xmean,clevel,tl,tu,ifail)

!     Display CI
      Write (nout,*) '95 percent Confidence Interval for the estimate'
      Write (nout,99998) '(', tl, ' ,', tu, ' )'
      Write (nout,*)

!     Calculate 99% confidence intervale
      clevel = 0.99E0_nag_wp
      ifail = 0
      Call g07abf(n,xmean,clevel,tl,tu,ifail)

!     Display CI
      Write (nout,*) '99 percent Confidence Interval for the estimate'
      Write (nout,99998) '(', tl, ' ,', tu, ' )'

99999 Format (1X,A,F7.4)
99998 Format (6X,A,F7.4,A,F7.4,A)
    End Program g07abfe
```

9.2 Program Data

```
G07ABF Example Program Data
0 3
1 17
2 26
3 16
4 18
5 9
6 3
7 5
8 0
9 1
10 0
```

9.3 Program Results

```
G07ABF Example Program Results

The point estimate of the Poisson parameter = 3.0204

95 percent Confidence Interval for the estimate
     ( 2.6861 , 3.3848 )

99 percent Confidence Interval for the estimate
     ( 2.5874 , 3.5027 )
```