Algorithm AS 91: The Percentage Points of the χ² Distribution
Author(s): D. J. Best and D. E. Roberts
Source: *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 24, No. 3 (1975), pp. 385-388
Published by: Blackwell Publishing for the Royal Statistical Society
Stable URL: http://www.jstor.org/stable/2347113
Accessed: 19/09/2010 10:43

Algorithm AS 91

The Percentage Points of the χ^2 Distribution

By D. J. Best and D. E. Roberts

C.S.I.R.O. Division of Mathematics and Statistics, North Ryde, Australia

Keywords: CHISQUARED DISTRIBUTION; PERCENTAGE POINTS; TAYLOR SERIES

Language

ISO Fortran

Description and Purpose

Given a value P of the lower tail area of the χ^2 distribution with ν degrees of freedom, the subroutine computes the corresponding percentage point z. Thus

$$P = \int_0^z \phi(u)\,du, \tag{1}$$

where

$$\phi(u) = 2^{-\frac{1}{2}\nu}\{\Gamma(\tfrac{1}{2}\nu)\}^{-1}\exp(-\tfrac{1}{2}u)\,u^{\frac{1}{2}\nu-1}, \quad \nu>0.$$

The subroutine is written so that z may be calculated (for ν not necessarily integral) as exactly as the user's computer allows. Thus our subroutine is both more general and more accurate than A.C.M. Algorithm 451 (Goldstein, 1973).

Numerical Method

z is found from the Taylor series expansion (Hill and Davis, 1968)

$$z = z_0 + \sum_r c_r(z_0)\{E/\phi(z_0)\}^r (r!)^{-1}, \tag{2}$$

where z_0 is a suitable starting approximation,

$$c_1(u) = 1, \quad c_{r+1}(u) = (r\psi + d/du)\,c_r(u),$$

$$E = P - \int_0^{z_0}\phi(u)\,du \quad \text{and} \quad \psi = \tfrac{1}{2} - (\tfrac{1}{2}\nu - 1)\,u^{-1}.$$

For many P, ν values the Wilson-Hilferty approximation (Kendall and Stuart, p. 372, 1969) can be used for z_0, viz.

$$z_{01} = \nu\{x(2/9\nu)^{\frac{1}{2}} + 1 - (2/9\nu)\}^3,$$

where x is the lower $100P\%$ point of the standard Normal distribution. However, better starting approximations are necessary in the three limiting cases, $P \to 0$, $P \to 1$ and $\nu \to 0$.

(i) $P \to 0$ (small z): Equation (1) can be simplified to give

$$z_{02} = \{P\nu 2^{\frac{1}{2}\nu-1}\,\Gamma(\tfrac{1}{2}\nu)\}^{2/\nu}.$$

z_{02} is better than z_{01} for $\nu < -1{\cdot}24 \ln P$. This criterion ensures that replacing $\exp(-\frac{1}{2}u)$ by 1 in (1) is in error by less than 10 per cent for $\nu \to 0$. For the special case $z_{02} < 2 \times 10^{-6}$, $z = z_{02}$ gives at least six significant figure accuracy.

(ii) $P \to 1$ (large z): Equation (1) can be simplified to give

$$z \doteq -2[\ln(1-P) - (\tfrac{1}{2}\nu - 1)\ln(\tfrac{1}{2}z) + \ln\{\Gamma(\tfrac{1}{2}\nu)\}].$$

For $z_{01} > 2{\cdot}2\nu + 6$ a better starting approximation than z_{01} is found to be

$$z_{03} = -2[\ln(1-P) - (\tfrac{1}{2}\nu - 1)\ln(\tfrac{1}{2}z_{01}) + \ln\{\Gamma(\tfrac{1}{2}\nu)\}].$$

(iii) $\nu \to 0$: For the special case $\nu \leqslant 0{\cdot}32$, P is expressed in terms of an approximation (Hastings, 1955) to the exponential integral and z_{04} found by Newton–Raphson iteration.

Structure

FUNCTION PPCHI2(P, V, G, IFAULT)

Formal parameters

P	Real	input:	value of lower tail area
V	Real	input:	degrees of freedom parameter
G	Real	input:	the natural logarithm of $\Gamma(\frac{1}{2}\nu)$
IFAULT	Integer	output:	a fault indicator, equal to: 1 if $P > 0{\cdot}999998$ or $P < 0{\cdot}000002$ 2 if $\nu \leqslant 0{\cdot}0$ 3 if the fault indicator of *FUNCTION GAMAIN* is greater than zero 0 otherwise

If a fault is detected *PPCHI2* is set equal to $-1{\cdot}0$.

Auxiliary algorithms

The following auxiliary subroutines are called:

FUNCTION GAMAIN(X, P, G, IFAULT)—Algorithm AS 32 (Bhattacharjee, 1970) and

FUNCTION GAUINV(P, IFAULT)—Algorithm AS 70 (Odeh and Evans, 1974).

For the natural logarithm of $\Gamma(\frac{1}{2}\nu)$ any standard algorithm, such as A.C.M. Algorithm 291, may be used.

Accuracy

If the appropriate starting approximation is used with seven terms in (2) only one evaluation is necessary to give at least six significant figures except for the small region $5{\cdot}6 < \nu < -4{\cdot}07 \ln(P) + 12{\cdot}21$ where two evaluations are necessary.

If more than six significant figures are required the *DATA* statement should be changed to alter E appropriately. When this is done more iterations of the first seven terms of (2) are performed as necessary.

References

Bhattacharjee, G. P. (1970). The incomplete gamma integral. *Appl. Statist.*, **19**, 285–287.

Goldstein, R. B. (1973). Algorithm 451: Chi-square quantiles. *Commun. Ass. Comput. Mach.*, **16**, 483–485.

HASTINGS, C., JR (1955). *Approximations for Digital Computers*. Princeton: University Press.
HILL, G. W. and DAVIS, A. W. (1968). Generalized asymptotic expansions of Cornish–Fisher type. *Ann. Math. Statist.*, **39**, 1264–1273.
KENDALL, M. G. and STUART, A. (1969). *The Advanced Theory of Statistics*, Vol. 1. London: Griffin.
ODEH, R. E. and EVANS, J. O. (1974). The percentage points of the normal distribution. *Appl. Statist.*, **22**, 96–97.
PIKE, M. C. and HILL, I. D. (1966). Algorithm 291: Logarithm of the gamma function. *Commun. Ass. Comput. Mach.*, **9**, 684.

```
      FUNCTION PPCHI2(P,V,G,IFAULT)
C
C        ALGORITHM AS 91  APPL. STATIST. (1975) VOL.24, NO.3
C
C        TO EVALUATE THE PERCENTAGE POINTS OF THE CHI-SQUARED
C        PROBABILITY DISTRIBUTION FUNCTION.
C        P MUST LIE IN THE RANGE 0.000002 TO 0.999998, V MUST BE POSITIVE,
C        G MUST BE SUPPLIED AND SHOULD BE EQUAL TO LN(GAMMA(V/2.0))
C
      DATA E, AA /0.5E-6, 0.6931471805/
C
C        AFTER DEFINING ACCURACY AND LN(2), TEST ARGUMENTS AND INITIALIZE
C
      PPCHI2 = -1.0
      IFAULT = 1
      IF (P .LT. 0.000002 .OR. P .GT. 0.999998) RETURN
      IFAULT = 2
      IF (V .LE. 0.0) RETURN
      IFAULT = 0
      XX = 0.5 * V
      C = XX - 1.0
C
C        STARTING APPROXIMATION FOR SMALL CHI-SQUARED
C
      IF (V .GE. -1.24 * ALOG(P)) GOTO 1
      CH = (P * XX * EXP(G + XX * AA)) ** (1.0 / XX)
      IF (CH - E) 6, 4, 4
C
C        STARTING APPROXIMATION FOR V LESS THAN OR EQUAL TO 0.32
C
    1 IF (V .GT. 0.32) GOTO 3
      CH = 0.4
      A = ALOG(1.0 - P)
    2 Q = CH
      P1 = 1.0 + CH * (4.67 + CH)
      P2 = CH * (6.73 + CH * (6.66 + CH))
      T = -0.5 + (4.67 + 2.0 * CH) / P1 -
     *    (6.73 + CH * (13.32 + 3.0 * CH)) / P2
      CH = CH - (1.0 - EXP(A + G + 0.5 * CH + C * AA) * P2 / P1) / T
      IF (ABS(Q / CH - 1.0) - 0.01) 4, 4, 2
C
C        CALL TO ALGORITHM AS 70 - NOTE THAT P HAS BEEN TESTED ABOVE
C
    3 X = GAUINV(P,IF1)
C
C        STARTING APPROXIMATION USING WILSON AND HILFERTY ESTIMATE
C
      P1 = 0.222222 / V
      CH = V * (X * SQRT(P1) + 1.0 - P1) ** 3
C
C        STARTING APPROXIMATION FOR P TENDING TO 1
C
      IF (CH .GT. 2.2 * V + 6.0)
     *   CH = -2.0 * (ALOG(1.0 - P) - C * ALOG(0.5 * CH) + G)
```

```
C         CALL TO ALGORITHM AS 32 AND CALCULATION OF SEVEN TERM
C         TAYLOR SERIES
C
    4 Q = CH
      P1 = 0.5 * CH
      P2 = P - GAMAIN(P1, XX, G, IF1)
      IF (IF1 .EQ. 0) GOTO 5
      IFAULT = 3
      RETURN
    5 T = P2 * EXP(XX * AA + G + P1 - C * ALOG(CH))
      B = T / CH
      A = 0.5 * T - B * C
      S1 = (210.0+A*(140.0+A*(105.0+A*(84.0+A*(70.0+60.0*A))))) / 420.0
      S2 = (420.0+A*(735.0+A*(966.0+A*(1141.0+1278.0*A)))) / 2520.0
      S3 = (210.0 + A * (462.0 + A * (707.0 + 932.0 * A))) / 2520.0
      S4 =(252.0+A*(672.0+1182.0*A)+C*(294.0+A*(889.0+1740.0*A)))/5040.0
      S5 = (84.0 + 264.0 * A + C * (175.0 + 606.0 * A)) / 2520.0
      S6 = (120.0 + C * (346.0 + 127.0 * C)) / 5040.0
      CH = CH+T*(1.0+0.5*T*S1-B*C*(S1-B*(S2-B*(S3-B*(S4-B*(S5-B*S6))))))
      IF (ABS(Q / CH - 1.0) .GT. E) GOTO 4
C
    6 PPCHI2 = CH
      RETURN
      END
```

Algorithm AS 92

The Sample Size for a Distribution-free Tolerance Interval

By P. Brooker and M. J. P. Selby

Civil Service Department *Peat, Marwick, Mitchell & Co.*

Keywords: SAMPLE SIZE; DISTRIBUTION FREE; SUCCESSIVE APPROXIMATION

Language

ISO Fortran

Purpose

If a sample of n independent and identically distributed random variables is chosen, this algorithm determines how large n must be, in order that, with probability at least β, the n chosen random variables will span a range which includes a proportion α or more of their parent distribution function.

Background

If l and u are the smallest and largest of n independent random variables, each with continuous c.d.f. $F(x)$, then the probability that the sample covers at least a proportion α of the parent distribution is the probability that

$$F(u)-F(l) \geqslant \alpha$$

which can be shown (Kendall and Stuart, 1967) to be $1-\phi(n)$ where $\phi(n) = n\,\alpha^{n-1} - (n-1)\,\alpha^n$, and to be independent of F. The smallest value of n, such that with