5-1967

Numerical Approximations to the Cumulative Chi-Square Distribution, the Cumulative t-Distribution and the Cumulative F-Distribution for Digital Computers

Grace Yuan-Chuen Wang

Utah State University

NUMERICAL APPROXIMATIONS TO THE CUMULATIVE CHI-SQUARE DISTRIBUTION, THE CUMULATIVE t-DISTRIBUTION AND THE CUMULATIVE F-DISTRIBUTION FOR DIGITAL COMPUTERS

by

Grace Yuan-Chuen Wang

A report submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Applied Statistics

Plan B

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1967

ACKNOWLEDGMENTS

The writer wishes to express her deepest gratitude to her major professor and report director, Dr. Rex L. Hurst, for his valuable suggestions and unfailing advice. She is also grateful to Dr. Richard Hansen and Dr. David White for their advice.

Grace Yuan-Chuen Wang

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

CHAPTER I

INTRODUCTION

There are good tables of the frequently used cumulative frequency distributions. These tables have some limitations with respect to the number of percentage points that are available. The main drawback in computer usage of these tables is that large amounts of storage and elaborate search and interpolation techniques are necessary for their use.

It is the purpose of this study to present associated numerical methods for digital computer which are satisfactorily accurate and which are reasonably economical in both time and machine memory capacity. To carry out this objective the following procedures were used:

1. A review of literature on numerical approximations—both texts and articles from statistical journals and computer science publications.

2. Writing test programs in Fortran for all the associated methods which can be obtained.

3. Checking the answers obtained by numerical approximation with the known answers in the table in order to determine usefulness of the numerical method.

4. Writing Fortran subprograms to evaluate those integrals by using the most accurate methods according to the experimental results.

CHAPTER II

NUMERICAL APPROXIMATION TO THE CUMULATIVE CHI-SQUARE DISTRIBUTION

2.1 Definitions

The probability that the variable $\frac{vs^2}{\sigma^2} = t$ is less than or equal to Chi-square is given by

$$P_m(\chi^2) = \left[2^{\frac{m}{2}} \Gamma(\tfrac{m}{2})\right]^{-1} \int_0^{\chi^2} t^{\frac{m}{2}-1} e^{\frac{-t}{2}} dt \qquad (2.1.1)$$

The constant m is called the number of degrees of freedom and

$\Gamma(\frac{m}{2}) = (\frac{m}{2} - 1)!$

The probability that the variable greater than or equal to Chi-square is given by,

$$Q_m(\chi^2) = \left[2^{\frac{m}{2}} \Gamma(\tfrac{m}{2})\right]^{-1} \int_{\chi^2}^{\infty} t^{\frac{m}{2}-1} e^{\frac{-t}{2}} dt \qquad (2.1.2)$$

The probability of the Chi-square in the interval $(0 \leqslant \chi^2 \leqslant \infty)$ is equal 1, i.e.,

$$\left[2^{\frac{m}{2}} \Gamma(\tfrac{m}{2})\right]^{-1} \int_0^{\infty} t^{\frac{m}{2}-1} e^{-\frac{t}{2}} dt = 1$$

and hence $Q_m(\chi^2) = 1 - P_m(\chi^2)$ (2.1.3)

2.2 Series Expansion

The series expansion for $Q_m(\chi^2)$ has been given by Elderton (1902).

$$\chi = \sqrt{\text{Chi-square}}$$

If m is even, $Q_m(\chi^2)$ will be calculated from

$$Q_m(\chi^2) = e^{\frac{1}{2}\chi^2}\left(1 + \frac{\chi^2}{2} + \frac{\chi^4}{2\cdot 4} + \frac{\chi^6}{2\cdot 4\cdot 6} + \cdots + \frac{\chi^{m-2}}{2\cdot 4\cdot 6\cdots(m-2)}\right) \tag{2.2.1}$$

If m is odd and $m > 1$, $Q_m(\chi^2)$ will be calculated from

$$Q_m(\chi^2) = {}^*\sqrt{\frac{2}{\pi}}\int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}\, d\chi + \sqrt{\frac{2}{\pi}}\, e^{-\frac{1}{2}\chi^2}\left(\frac{\chi}{1} + \frac{\chi^3}{1\cdot 3}\right.$$

$$\left. + \frac{\chi^5}{1\cdot 3\cdot 5} + \cdots + \frac{\chi^{m-2}}{1\cdot 3\cdot 5\cdots(m-2)}\right) \tag{2.2.2}$$

2.3 Approximations

The χ^2-distribution converges slowly to the normal distribution. The simplest transformation has been given by R. A. Fisher (1925) who has demonstrated that for $m > 30$. The quantity $\sqrt{2\chi^2} - \sqrt{2m-1}$ is nearly normally distributed with zero mean and unit variance, i.e.,

---

*Approximation to $\sqrt{\frac{2}{\pi}}\int_{\chi}^{\infty} e^{\frac{1}{2}}\, d\chi$ (see Appendix).

$u \sim (0,1)$

$$u \doteq \sqrt{2\chi^2} - \sqrt{2m-1} \qquad (2.3.1)$$

The approximation (2.3.1) is not very accurate. A more accurate approximation is given by Wilson and Hilferty (1931)

$$*u \doteq \frac{\left(\frac{\chi^2}{m}\right)^{\frac{1}{3}} - \left(1 - \frac{2}{9m}\right)}{\sqrt{\frac{2}{9m}}} \quad \text{for } (m > 30) \qquad (2.3.2)$$

According to the writer's experimental results, both the approximation (2.3.1) and the approximation (2.3.2) are not accurate for $\chi^2 \leqslant m$. A new approximation to $P_m(\chi^2)$ is given by Hastings (1954). It holds to closer than 0.0003 for $0 \leqslant \chi^2 \leqslant m$ and $2 \leqslant m \leqslant \infty$

$$P_m(\chi^2) = \frac{A}{(1 + a_1 y + a_2 y^2 + a_3 y^3 + a_4 y^4)^4} \qquad (2.3.3)$$

$$y = \sqrt{\frac{m}{2}} \ \ln\left(\frac{m}{\chi^2}\right)$$

$$A = 0.5 + 0.1323\left(\frac{2}{m}\right)^{\frac{1}{2}} + 0.0036\left(\frac{2}{m}\right) - 0.0038\left(\frac{2}{m}\right)^{\frac{3}{2}}$$

$$a_1 = 0.1968 - 0.0452\left(\frac{2}{m}\right)^{\frac{1}{2}} - 0.0128\left(\frac{2}{m}\right) - 0.0168\left(\frac{2}{m}\right)^{\frac{3}{2}}$$

---

*Approximation to u-distribution (see Appendix).

$$a_2 = 0.1152 - 0.0990 \left(\frac{2}{m}\right)^{\frac{1}{2}} + 0.0539 \left(\frac{2}{m}\right) - 0.0168 \left(\frac{2}{m}\right)^{\frac{3}{2}}$$

$$a_3 = 0.0004 + 0.0442 \left(\frac{2}{m}\right)^{\frac{1}{2}} - 0.0866 \left(\frac{2}{m}\right) + 0.0398 \left(\frac{2}{m}\right)^{\frac{3}{2}}$$

$$a_4 = 0.0195 - 0.0629 \left(\frac{2}{m}\right)^{\frac{1}{2}} + 0.0708 \left(\frac{2}{m}\right) - 0.0269 \left(\frac{2}{m}\right)^{\frac{3}{2}}$$

2.4 Recurrence Relation

The recurrence relation has been given by Zelen and Severo (1965)

$$Q_{m+2}(\chi^2) = Q_m(\chi^2) + \frac{\left(\frac{\chi^2}{m}\right)^{\frac{m}{2}} e^{-\frac{\chi^2}{2}}}{\Gamma\left(\frac{m}{2} + 1\right)} \qquad (2.4.1)$$

Example:

Let m = 1

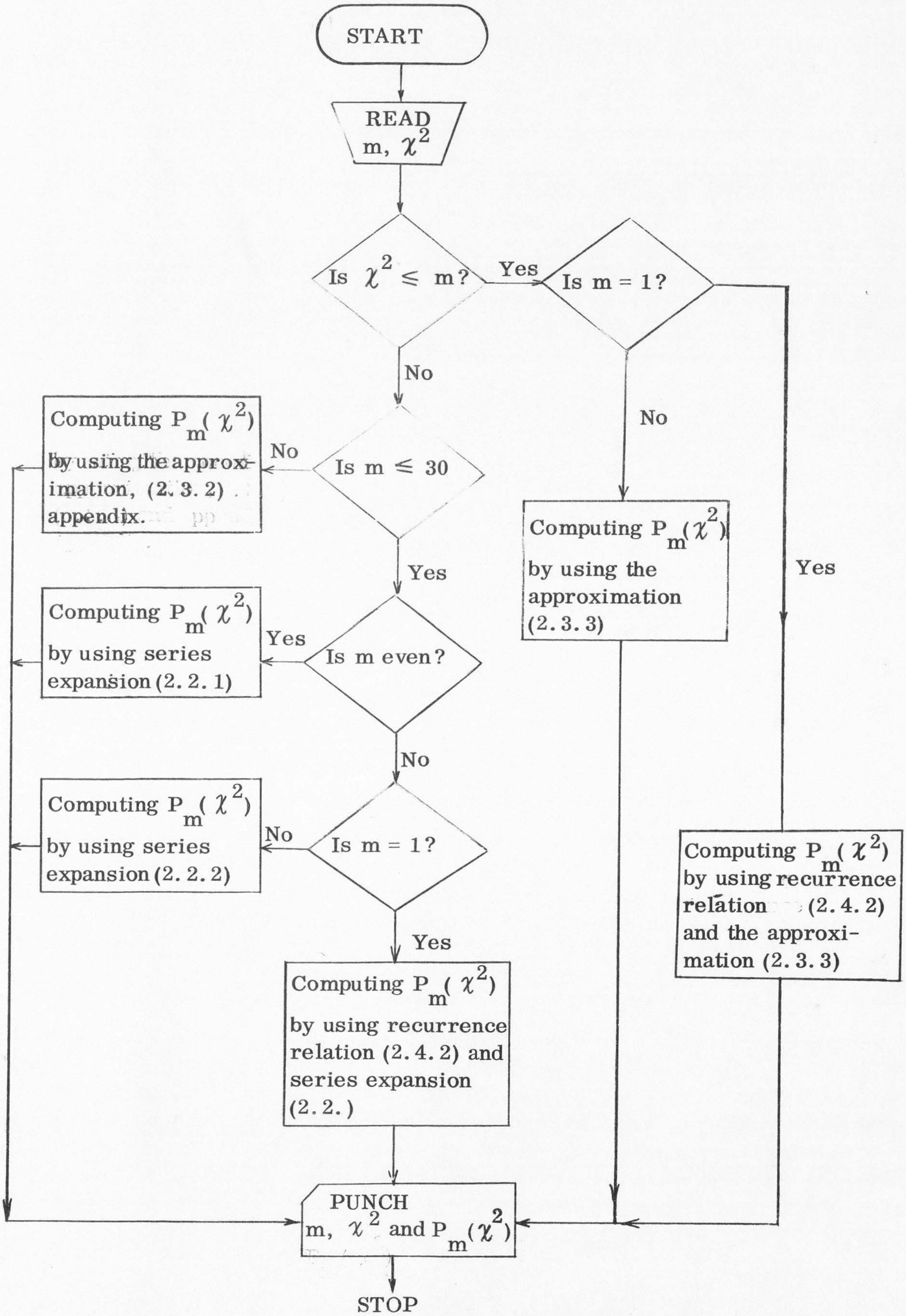
$$Q_1(\chi^2) = Q_3(\chi^2) - \frac{\left(\frac{\chi^2}{2}\right)^{\frac{1}{2}} e^{-\frac{\chi^2}{2}}}{\Gamma\left(\frac{1}{2} + 1\right)}$$

$$\Gamma\left(\frac{1}{2} + 1\right) = \frac{1}{2}! = \frac{1}{2}\left(-\frac{1}{2}\right)! = \frac{1}{2}\sqrt{\pi} = 0.886225$$

Hence, $$Q_1(\chi^2) = Q_3(\chi^2) - \frac{\left(\frac{\chi^2}{2}\right)^{\frac{1}{2}} e^{-\frac{\chi^2}{2}}}{0.886225} \qquad (2.4.2)$$

2.5 The Calculation Procedures for $P_m(\chi^2)$ when χ^2 and m are given

No one computing procedure or approximation is satisfactory for all region and degrees of freedom. A control program is necessary to isolate region. A different procedure may be used in each region to meet requirement of accuracy and computing time. Procedures are illustrated by the following flow chart.


START
READ
m, χ^2
Is $\chi^2 \leq$ m?
Yes
Is m = 1?
No
No
Computing $P_m(\chi^2)$ by using the approximation, (2. 3. 2) appendix.
No
Is m $\leq$ 30
Computing $P_m(\chi^2)$ by using the approximation (2. 3. 3)
Yes
Yes
Computing $P_m(\chi^2)$ by using series expansion (2. 2. 1)
Yes
Is m even?
No
Computing $P_m(\chi^2)$ by using series expansion (2. 2. 2)
No
Is m = 1?
Computing $P_m(\chi^2)$ by using recurrence relation (2. 4. 2) and the approximation (2. 3. 3)
Yes
Computing $P_m(\chi^2)$ by using recurrence relation (2. 4. 2) and series expansion (2. 2.)
PUNCH
m, χ^2 and $P_m(\chi^2)$
STOP

2.6 Test Program

The following is a listing of a Fortran IV program which computes the probability when. χ^2, m and the tabular values are read in. A difference in probability values is computed in order to determine whether or not the computation by the calculation procedures presented in the previous section is worthwhile.

The output for the program gives the probability p obtained by machine computation and the difference (comparing the computed answer p with known answer tabular p). The output shows the absolute error is between 0.00000 and 0.00095, hence, the composite procedure is quite accurate.

Fortran program for the cumulative
Chi-square distribution

```
C        THE IS A TEST PROGRAM TO CHECK THE ANSWERS OBTAINED BY
C        MACHINE COMPUTATION FOR NUMERICAL APPROXIMATIONS WITH
C        THE KNOWN ANSWERS IN THE TABLE
C        BY GRACE YUAN-CHUEN WANG, UTAH STATE UNIVERSITY
         WRITE (3,10)
      10 FORMAT (1H1,7X,6HOUTPUT//)
         WRITE(3,11)
      11 FORMAT(5X4HD.F.,4X,10HCHI-SQUARE,4X,5HEST P,7X,5HTAB P,
        14X,4HDIFF//)
     100 READ(2,20) V,CHISQ,TRUEP
      20 FORMAT (F5.0,2F10.5)
         CALL PROBC(V,CHISQ,ESTP)
         ERROR=TRUEP-ESTP
         M=V
```

```
      WRITE(3,12) M, CHISQ, ESTP, TRUEP, ERROR
   12 FORMAT(4X,I5,4(2X,F10.5) )
      GO TO 100
      END

      SUBROUTE PROBC(V, CHISQ, P)
C     THE CUMULATIVE CHI-SQUARE DISTRIBUTION
C     COMPUTING PROBABILITY WHEN DEGREE OF FREEDOM AND
C     CHI-SQUARE ARE READ IN
      IV=V
      W=.5*CHISQ
      IF(CHISQ-V) 102,102,103
C     COMPUTING P WHEN CHI-SQUARE IS LESS THAN THE NUMBER OF
C     DEGREES OF FREEDOM
  102 IF(IV-1) 10,10,12
   10 AM=V+2.
      GO TO 14
   12 AM=V
   14 ANU=SQRT (AM/2.)*ALOG(AM/CHISQ)
      C1=SQRT (2./AM)
      C2=2./AM
      C3=C1*C2
      A=.5+.1323*C1+.0036*C2-.0038*C3
      A1=.1968-.0452*C1-.0128*C2+.0067*C3
      A2=.1152-.0990*C1+.0539*C2-.0168*C3
      A3=.0004+.0442*C1-.0866*C2+.0398*C3
      A4=.0195-.0629*C1+.0708*C2-.0269*C3
       P=A/( (1.+A1*ANU+A2*(ANU*ANU)+A3*(ANU**3)+A4*(ANU**4) )**4)
      IF(IV-1)16,16,99
   16  P= P+(SQRT  (W)*(1./EXP (W) ) )/.886225
      GO TO 99
  103 IF(IV-30) 104,104,105
C     COMPUTING P WHEN CHI-SQUARE IS LARGE THAN THE NUMBER
C     OF DEGREES OF FREEDOM AND THE NUMBER OF DEGREES OF
C     FREEDOM IS LESS THAN 30
  104 X=SQRT (CHISQ)
      IF( (-1)**IV) 106,106,107
  106 IF (IV-1)71,71,72
   71 M=IV+2
      GO TO 70
   72 M=IV
   70 SUME=0.
      PE=1.
      N=M-2
      DO 77 J=l,N,2
      PJ=J
      PE=PE*PJ
```

```
   77 SUME=SUME+X**J/PE
      Y=SQRT (0.5*CHISQ)
      S1=.0706230784*Y
      S2=.0422820123*Y*Y
      S3=.0092705272*Y**3
      S4=.0001520143*Y**4
      S5=.0002765672*Y**5
      S6=.0000430638*Y**6
      TEG=1./(1.+S1+S2+S3+S4+S5+S6)**16
      P=1.-TEG-SQRT (2./3.141593)*(1./EXP(W))*SUME
      IF(IV-1) 76,76,99
   76 P=P+(SQRT (W)*(1.0/EXP (W) ) )/0.886225
      GO TO 99
  107 IF(IV-2) 61,61,60
   61 P=1.-1./EXP (W)
      GO TO 99
   60 SUMO=0.
      PO=1.
      K=IV-2
      DO 66 I=2,K,2
      PI=I
      PO=PO*PI
   66 SUMO=SUMO+X**I/PO
      P=1.-(1./EXP (W) )*(1.+SUMO)
      GO TO 99
C     COMPUTING P WHEN CHI-SQUARE IS LARGE THAN THE NUMBER OF
C     DEGREES OF FREEDOM AND THE NUMBER OF DEGREES OF FREEDOM
C     IS LARGE THAN 30
  105 A=2./(9.*V)
      B=(CHISQ/V)**(1./3.)
      X=(B-(1.-A) )/SQRT (A)
      ZX=1./(2.5052367+1.2831204*X*X+.2264718*X**4+.1306469*X**6
     1 -.0202490*X**8+.0039132*X**10)
      C=1./(1.+.2316419*X)
      P=1.-ZX*(.319381530*C-.356563782*C*C+1.781477937*C**3
     1-1.821255978*C**4+1.330274429*C**5)
   99 RETURN
      END
```

OUTPUT

D. F.	CHI-SQUARE	EST P	TAB P	DIFF
120	83.85000	.00512	.00500	-.00012
40	22.16000	.00999	.01000	.00001
60	37.48000	.00999	.01000	.00001
120	86.92000	.00999	.01000	.00001
40	24.43000	.02485	.02500	.00015
60	40.48000	.02483	.02500	.00017
120	91.58000	.02485	.02500	.00015
1	.01580	.10003	.10000	-.00003
2	.21070	.10012	.10000	-.00012
3	.58400	.09987	.10000	.00013
4	1.06400	.10016	.10000	-.00016
5	1.61000	.10013	.10000	-.00013
6	2.20000	.09978	.10000	.00022
10	4.87000	.10049	.10000	-.00049
40	29.05000	.10006	.10000	-.00006
120	100.62000	.10002	.10000	-.00002
1	.00390	.04980	.05000	.00020
2	.10260	.05022	.05000	-.00022
3	.35200	.04988	.05000	.00012
4	.71100	.05008	.05000	-.00008
5	1.15000	.05054	.05000	-.00054
6	1.64000	.05049	.05000	-.00049
7	2.17000	.05031	.05000	-.00031
8	2.73000	.04999	.05000	.00001
9	3.33000	.05035	.05000	-.00035
10	3.94000	.05009	.05000	-.00009
11	4.57000	.04989	.05000	.00011
12	5.23000	.05021	.05000	-.00021
14	6.57000	.05002	.05000	-.00002
15	7.26000	.05000	.05000	.00000
16	7.96000	.04996	.05000	.00004
18	9.39000	.04998	.05000	.00002
20	10.85000	.04996	.05000	.00004
24	13.85000	.04998	.05000	.00002
30	18.49000	.04987	.05000	.00013
40	26.51000	.04990	.05000	.00010
60	43.19000	.04988	.05000	.00012
120	95.70000	.04979	.05000	.00021
1	.45500	.49993	.50000	.00007
1	.06420	.19996	.20000	.00004
1	.14800	.29940	.30000	.00060
1	.27500	.39984	.40000	.00016

1	.70800	.59990	.60000	.00010
1	1.07000	.69905	.70000	.00095
1	1.64000	.79967	.80000	.00033
1	2.71000	.90028	.90000	-.00028
40	51.81000	.90029	.90000	-.00029
60	74.40000	.90020	.90000	-.00020
120	140.23000	.90008	.90000	-.00008
3	11.34000	.98998	.99000	.00002
6	16.81000	.98999	.99000	.00001
11	24.73000	.99002	.99000	-.00002
16	32.00000	.99000	.99000	-.00000
24	42.98000	.99000	.99000	-.00000
30	50.89000	.98999	.99000	.00001
40	63.69000	.98997	.99000	.00003
60	88.38000	.98999	.99000	.00001
1	5.02000	.97494	.97500	.00006
1	5.02000	.97494	.97500	.00006
1	6.63000	.98997	.99000	.00003
1	6.63000	.98997	.99000	.00003
1	7.88000	.99500	.99500	-.00000
1	10.80000	.99899	.99900	.00001
1	12.10000	.99950	.99950	.00000
1	3.84000	.94996	.95000	.00004
2	5.99000	.94996	.95000	.00004
3	7.81000	.94989	.95000	.00011
4	9.49000	.95005	.95000	-.00005
5	11.10000	.95057	.95000	-.00057
7	14.10000	.95057	.95000	-.00057
8	15.51000	.95004	.95000	-.00004
9	16.90000	.94969	.95000	.00031
11	19.70000	.95037	.95000	-.00037
14	23.68000	.94993	.95000	.00007
15	25.00000	.95006	.95000	-.00006
16	26.30000	.95005	.95000	-.00005
19	30.10000	.94946	.95000	.00054
21	32.70000	.95035	.95000	-.00035
23	35.20000	.95032	.95000	-.00032
25	37.70000	.95053	.95000	-.00053
26	38.90000	.95016	.95000	-.00016
27	40.10000	.94986	.95000	.00014
28	41.30000	.94960	.95000	.00040
29	42.60000	.95045	.95000	-.00045
30	43.77000	.94997	.95000	.00003
31	45.00000	.95025	.95000	-.00025
33	47.40000	.95010	.95000	-.00010
36	51.00000	.95011	.95000	-.00011

38	53.40000	.95024	.95000	-.00024
40	55.76000	.95010	.95000	-.00010
41	56.90000	.94970	.95000	.00030
43	59.30000	.95005	.95000	-.00005
46	62.80000	.94982	.95000	.00018
48	65.20000	.95033	.95000	-.00033
50	67.50000	.95004	.95000	-.00004
60	79.08000	.95005	.95000	-.00005
120	146.57000	.95007	.95000	-.00007
5	12.83000	.97497	.97500	.00003
24	39.36000	.97498	.97500	.00002
120	152.21000	.97489	.97500	.00011
3	11.34000	.98998	.99000	.00002
18	34.81000	.99001	.99000	-.00001
9	23.59000	.99500	.99500	-.00000
30	53.67000	.99500	.99500	.00000

CHAPTER III

NUMERICAL APPROXIMATION TO THE CUMULATIVE t-DISTRIBUTION

3.1 Definitions

The distribution of the variable $\frac{\theta - \mu_\theta}{s_\theta} = x$ in the interval $(-\infty \leq x \leq \infty)$ if given by

$$\frac{1}{\sqrt{\pi m}} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} (1 + \frac{x^2}{m})^{-\frac{m+1}{2}} \qquad (3.1.1)$$

The constant m is called the number of degrees of freedom

$$\Gamma(\frac{m+1}{2}) = (\frac{m-1}{2})! \text{ and } \Gamma(\frac{m}{2}) = (\frac{m-2}{2})!$$

The probability that the variable is less than or equal to t is given by

$$P_m(t) = \frac{1}{\sqrt{\pi m}} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \int_{-\infty}^{t} (1 + \frac{x^2}{m})^{-\frac{m+1}{2}} dx \qquad (3.1.2)$$

The probability that the variable falls within the interval $(-t \leq x \leq t)$ is given by

$$A_m(t) = \frac{1}{\sqrt{\pi m}} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \int_{-t}^{t} (1 + \frac{x^2}{m})^{-\frac{m+1}{2}} dx \qquad (3.1.3)$$

3.2 Series Expansion

The series expansion for $A_m(t)$ has been given by Zelen and Serevo (1965)

$$\theta = \arctan \frac{t}{\sqrt{m}}$$

If m is odd, $A_m(t)$ will be calculated from

$$A_m(t) = \begin{cases} \frac{2}{\pi}\left\{\theta + \sin\theta \left[(\operatorname{Cos}\theta + \frac{2}{3}\cos^3\theta + \cdots + \frac{2\cdot 4\cdots(m-3)}{1\cdot 3\cdots(m-2)}\cos^{m-2}\theta\right]\right\} & \text{for } (m > 1) \\ \frac{2}{\pi}\theta & \text{for } (m=1) \end{cases} \qquad (3.2.1)$$

If m is even, $A_m(t)$ will be calculated from

$$A_m(t) = \sin\theta\left[1 + \frac{1}{2}\cos^2\theta + \frac{1\cdot 3}{2\cdot 4}\cos^4\theta + \cdots + \frac{1\cdot 3\cdot 5\cdots(m-3)}{2\cdot 4\cdot 6\cdots(m-2)}\cos^{m-2}\theta\right] \qquad (3.2.2)$$

3.3 Approximation

The t-distribution is symmetrical about zero like the normal distribution. It converges to the normal distribution as degrees of freedom increase. An approximation to the t-distribution for large m has been given by Zelen and Serevo (1965). The quantity $\dfrac{t(1 - \frac{1}{4m})}{\sqrt{1 + \frac{t^2}{m}}}$ is nearly normally distributed with zero mean and unit variance, i.e.,

$u \sim (0, 1)$

$$*u \doteq \frac{t\left(1 - \frac{1}{4m}\right)}{\sqrt{1 + \frac{t^2}{2m}}} \qquad (3.3.1)$$

According to the writer's experimental results, the approximation (3.3.1) is accurate for $m > 17$ and $t \geqslant 0$.

3.4 Computing $P_m(t)$ For Negative Value of t

Since the t-ditribution is symmetrical about zero, for negative values of t, one uses the relation

$$P_m(-t) = 1 - P_m(t) \qquad (3.4.1)$$

For example

$$P_1(-1.00) = 1 - P_1(1.00)$$

$$= 1 - .75 = .25$$

$$P_m(t) = A_m(t) + P_m(-t)$$

$$= A_m(t) + 1 - P_m(t)$$

and hence

$$P_m(t) = \frac{1}{2}\left(1 + A_m(t)\right) \qquad (3.4.2)$$

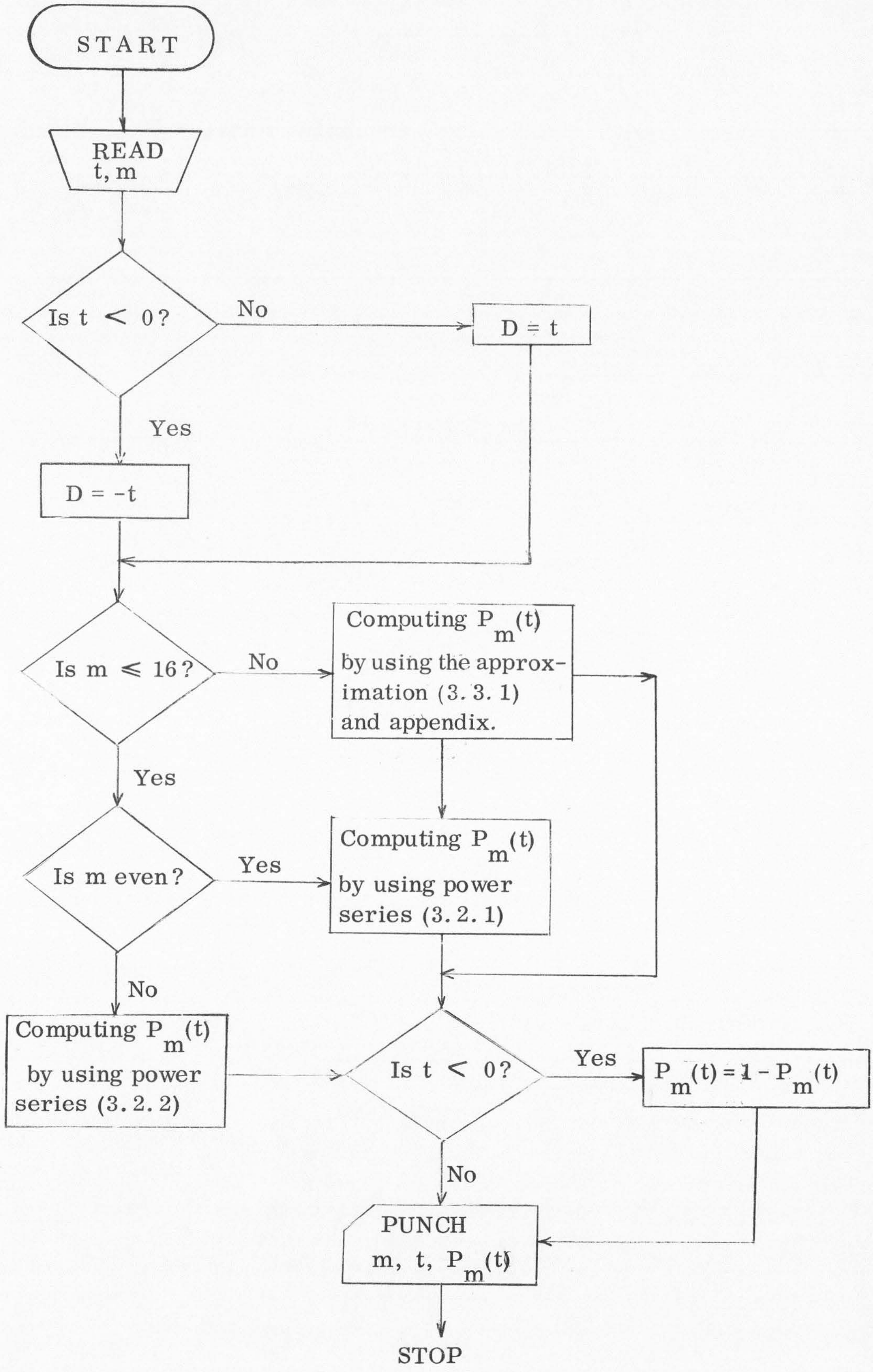
---

*Approximation to u-distribution (see Appendix).

3.5 The calculation Procedures for $P_m(t)$

When t and m are given

No one computing procedure or approximations is satisfactory for all region and degrees of freedom. A control program is necessary to isolate regions. A different procedure may be used in each region to meet requirements of accuracy and computing time. Procedures are illustrated by the following flow chart.


START
READ
t, m
Is t < 0?
No
D = t
Yes
D = -t
Is m ≤ 16?
No
Computing $P_m(t)$
by using the approximation (3. 3. 1)
and appendix.
Yes
Is m even?
Yes
Computing $P_m(t)$
by using power series (3. 2. 1)
No
Computing $P_m(t)$
by using power series (3. 2. 2)
Is t < 0?
Yes
$P_m(t) = 1 - P_m(t)$
No
PUNCH
m, t, $P_m(t)$
STOP

3.6 Test Program

Following is a listing of a Fortran IV program which computes the probability when t, m and the tabular values are read in. A difference in probability values is computed in order to determine whether or not computation by the calculation procedures presented in the previous section is worthwhile.

The output for the program gives probability obtained by machine computation and the difference (comparing the computed answer p with known answer tabular p). The output shows that absolute error is between 0.00000 and 0.00035, and hence the composite procedure is satisfactorily accurate.

Fortran program for the cumulative t-distribution

```
C       THE IS A TEST PROGRAM TO CHECK THE ANSWERS OBTAINED BY
C       MACHINE COMPUTATION FOR NUMERICAL APPROXIMATIONS WITH
C       THE KNOWN ANSWERS IN THE TABLE
C       BY GRACE YUAN-CHEUN WANG,UTAH STATE UNIVERSITY
        WRITE(3,10)
   10   FORMAT(1H1,7X,6HOUTPUT//)
        WRITE(3,11)
   11   FORMAT(5X,4HD.F.,6X,1HT,11X,5HEST P,7X,5HTAB P,6X,4HDIFF//)
  100   READ(2,20) V,T,TRUEP
   20   FORMAT(F5.0,2F10.5)
        CALL PROBT(V,T,ESTP)
        M=V
        ERROR=TRUEP-ESTP
        WRITE(3,12) M,T,ESTP,TRUEP,ERROR
   12   FORMAT(4X,I5,4(2X,F10.5))
        GO TO 100
        END
```

```
      SUBROUTINE PROBT(V,T,P)
C     THE CULUMATIVE T DISTRIBUTION
C     COMPUTING PROBABILITY WHEN V AND T ARE READ IN
C     V IS THE NUMBER OF DEGREES OF FREEDOM
      IV=V
      IF(T) 4,5,5
    4 D=-T
      GO TO 6
    5 D=T
    6 IF(IV-16) 102,102,103
C     COMPUTING P WHEN THE NUMBER OF DEGREES OF FREEDOM IS
C     LESS THAN 17
  102 B=ATAN (D/SQRT (V) )
      IF( (-1)**IV)23,24,24
   23 IF(IV-3) 25,26,27
   25 ATV=2.*B/3.1416
      GO TO 66
   26 ATV=2./3.1416*(B+SIN (B)*COS (B) )
      GO TO 66
   27 A=1.
      SUM=0.
      J=IV-2
      DO 20 N=3,J,2
      PN=N
      A=A*(PN-1.)/PN
   20 SUM=SUM+A*COS (B)**N
      ATV=2./3.1416*(B+SIN (B)*COS (B)+SIN  (B)*SUM)
      GO TO 66
   24 IF(IV-2)  28,29,28
   28 C=1.
      TOTAL=0.
      K=IV-2
      DO 21 N=2,K,2
      TN=N
      C=C*(TN-1.)/TN
   21 TOTAL=TOTAL+C*COS  (B)**N
      ATV=SIN  (B)*(1.+TOTAL)
      GO TO 66
   29 ATV=SIN (B)
   66 P=(1.+ATV)/2.
      GO TO 7
C     COMPUTING P WHEN THE NUMBER OF DEGREES OF FREEDOM
C     IS LARGE THAN 16
  103 X=D*(1.-1./(4.*V) )/SQRT  (1.+D*D/(2.*V) )
      ZX=1./(2.5052367+1.2831204*X*X+.2264718*X**4+.1306469*X**6
     1-.0202490*X**8+.0039132*X**10)
```

```
      C=1./(1.+.2316419*X)
      PX=1.-ZX*(.319381530*C-.356563782*C*C+1.781477937*C**3
     1-1.821255978*C**4+1.330274429*C**5)
      P=PX
    7 IF(T)8,99,99
    8 P=1.-P
   99 RETURN
      END
```

OUTPUT

D. F.	T	EST P	TAB P	DIFF
1	-12.70600	.02500	.02500	-.00000
4	-2.77600	.02501	.02500	-.00001
7	-2.36500	.02499	.02500	.00001
10	-2.22800	.02501	.02500	-.00001
12	-2.17900	.02499	.02500	.00001
15	-2.13100	.02502	.02500	-.00002
18	-2.10100	.02536	.02500	-.00036
23	-2.06900	.02524	.02500	-.00024
26	-2.05600	.02520	.02500	-.00020
30	-2.04200	.02520	.02500	-.00020
40	-2.02100	.02515	.02500	-.00015
120	-1.98000	.02510	.02500	-.00010
1	.32500	.60002	.60000	-.00002
4	.27100	.60010	.60000	-.00010
7	.26300	.59994	.60000	.00006
14	.25800	.59992	.60000	.00008
17	.25700	.59988	.60000	.00012
23	.25600	.59991	.60000	.00009
28	.25600	.60011	.60000	-.00011
40	.25500	.60001	.60000	-.00001
120	.25400	.60006	.60000	-.00006
1	6.31400	.95000	.95000	-.00000
2	2.92000	.95000	.95000	-.00000
3	2.35300	.94998	.95000	.00002
4	2.13200	.95001	.95000	-.00001
5	2.01500	,95000	.95000	.00000
6	1.94300	.94999	.95000	.00001
7	1.89500	.95003	.95000	-.00003
8	1.86000	.95003	.95000	-.00003
9	1.83300	.94999	.95000	.00001
10	1.81200	.94996	.95000	.00004

11	1.79600	.95001	.95000	-.00001
12	1.78200	.94998	.95000	.00002
13	1.77100	.95000	.95000	-.00000
14	1.76100	.94997	.95000	.00003
15	1.75300	.94999	.95000	.00001
16	1.74600	.95001	.95000	-.00001
17	1.74000	.94983	.95000	.00017
18	1.73400	.94982	.95000	.00018
19	1.72900	.94984	.95000	.00016
20	1.72500	.94989	.95000	.00011
21	1.82100	.94991	.95000	.00009
22	1.71700	.94989	.95000	.00011
23	1.71400	.94992	.95000	.00008
24	1.71100	.94993	.95000	.00007
25	1.70800	.94992	.95000	.00008
26	1.70600	.94998	.95000	.00002
27	1.70300	.94992	.95000	.00008
28	1.70100	.94994	.95000	.00006
29	1.69900	.94995	.95000	.00005
30	1.69700	.94994	.95000	.00006
40	1.68400	.95001	.95000	-.00001
60	1.67100	.95006	.95000	-.00006
120	1.65680	.95007	.95000	-.00007
1	31.82100	.99000	.99000	.00000
6	3.14300	.99000	.99000	-.00000
14	2.62400	.98999	.99000	.00001
18	2.55200	.98973	.99000	.00027
23	2.50000	.98985	.99000	.00015
28	2.46700	.98990	.99000	.00010
120	2.35800	.99002	.99000	-.00002
2	9.92500	.99500	.99500	-.00000

CHAPTER IV

NUMERICAL APPROXIMATION TO THE CUMULATIVE F-DISTRIBUTION

4.1 Definitions

The distribution of the variable $\frac{x_{1/m}}{x_{2/n}} = x$ in the interval $(0 \leq y < \infty)$ is given by

$$\frac{\left(\frac{m+n-2}{2}\right)!\; m^{\frac{m}{2}}\; n^{\frac{n}{2}}}{\left(\frac{m-2}{2}\right)!\left(\frac{n-2}{2}\right)!} \quad x^{\frac{(m-2)}{2}} (n+mx)^{-\frac{(m+n)}{2}}$$

The constant m is the number of degrees of freedom in the numerator of F and the constant n is the number of degrees of freedom in the denominator of F. x_1 and x_2 are distributed as Chi-square.

The probability that the variable is less than or equal to F is given by

$$P_{m,n}(F) = \frac{\left(\frac{m+n-2}{2}\right)!\quad m^{\frac{m}{2}}\; n^{\frac{n}{2}}}{\left(\frac{m-2}{2}\right)!\left(\frac{n-2}{2}\right)!} \int_0^F x^{\frac{(m-2)}{2}} (n+mx)^{-\frac{(m+n)}{2}}\, dx \tag{4.1.1}$$

4.2 Series Expansion

The series expansion for $P_{m,n}(F)$ has been given by Zelen and Serevo (1965)

$$x = \frac{n}{n + mF} \ ,$$

$$\theta = \text{arc tan} \sqrt{\frac{m}{n} F}$$

$$\phi = \text{arc tan} \sqrt{\frac{F}{n}}$$

If m is even, $P_{m,n}(F)$ will be calculated from

$$P_{m,n}(F) = 1 - x^{\frac{n}{2}} \left[1 + \frac{n}{2}(1 - x) + \frac{n(n+2)}{2 \cdot 4} (1 - x)^2 + \cdots + \frac{n(n+2) \cdots (n+m-4)}{2 \cdot 4 \cdots (m-2)} (1 - x)^{\frac{m-2}{2}} \right] \qquad (4.2.1)$$

If n is even, $P_{m,n}(F)$ will be calculated from

$$P_{m,n}(F) = (1 - x)^{\frac{m}{2}} \left[1 + \frac{m}{2} x + \frac{m(m+2)}{2 \cdot 4} x^2 + \cdots + \frac{m(m+2) \cdots (n+m-4)}{2 \cdot 4 \cdots (n-2)} x^{\frac{n-2}{2}} \right] \qquad (4.2.2)$$

If both m and n are odd and greater than 1, $P_{m,n}(F)$ will be calculated from

$$P_{m,n}(F) = A_n(t) - B(m,n) \qquad (4.2.3)$$

$$A_n(t) = \frac{2}{\pi} \left\{ \theta + \sin\theta \left[\cos\theta + \frac{2}{3} \cos^3\theta + \cdots + \frac{2.4 \cdots (n-3)}{1 \cdot 3 \cdots (n-2)} \cos^{n-2}\theta \right] \right\}$$

$$B(m,n) = \frac{2}{\sqrt{\pi}} \frac{\left(\frac{n-1}{2}\right)!}{\left(\frac{n-2}{2}\right)!} \sin\theta \cos^n\theta \left[1 + \frac{n+1}{3} \sin^2\theta + \cdots + \frac{(n+1)(n+3) \cdots (m+n-4)}{3 \cdot 5 \cdots (m-2)} \sin^{m-3}\theta \right]$$

If n = 1, m is odd and greater than 1, $P_{m,n}(F) = \frac{2\theta}{\pi} - B(m,n)$ (4.2.4)

If m = 1 and n is odd, $P_{m,n}(F)$ will be calculated from

$$P_{m,n}(F) = \begin{cases} \frac{2}{\pi}\left\{\phi + \sin\phi\left[\cos\phi + \frac{2}{3}\cos^3\phi + \cdots + \frac{2\cdot 4\cdots(n-3)}{1\cdot 3\cdots(n-2)}\cos^{n-2}\phi\right]\right\} & \text{for } (n > 1) \\ \frac{\pi}{2}\,\theta & \text{for } (n = 1) \end{cases} \quad (4.2.5)$$

4.3 Approximation

A quite accurate approximation for large values of m and n has been given by Zelen and Serevo (1965). The quantity $\frac{F^{\frac{1}{3}}(1 - \frac{2}{9n}) - (1 - \frac{2}{9m})}{\sqrt{\frac{2}{9m} + F^{\frac{2}{3}}(\frac{2}{9n})}}$

is nearly normally distributed with zero mean and unit variance, i.e.,

$$u \sim (0.1)$$

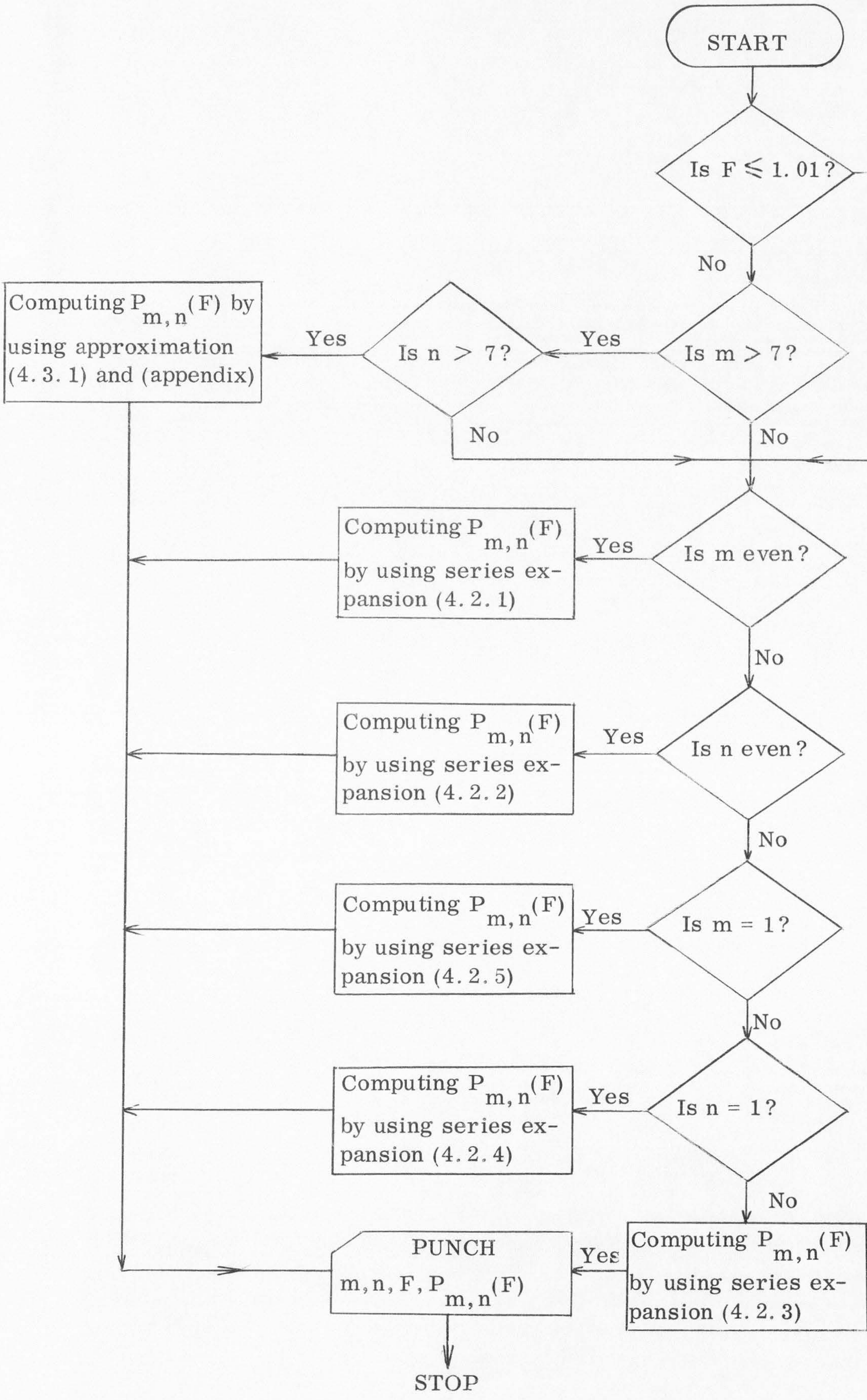
$$^{*}u \doteq \frac{F^{\frac{1}{3}}(1 - \frac{2}{9n}) - (1 - \frac{2}{9m})}{\sqrt{\frac{2}{9m} + F^{\frac{2}{3}}(\frac{2}{9n})}}$$

*Approximation to u-distribution (see Appendix).

According to the writer's experimental result, the approximation (4.3.1) is accurate for $m > 7$, $n > 7$ and $F > 1.01$.

4.4 The Calculation Procedure For $P_{m,n}(F)$ When m, n and F Are Given

No one computing procedure or approximation is satisfactory for all regions and degrees of freedom. A control program is necessary to isolate regions. A different procedure may be used in each region to meet requirements of accuracy and computing time. Procedures are illustrated by the following flow chart.


START
Is F ≤ 1.01?
Yes
No
Is m > 7?
Yes
Is n > 7?
Yes
Computing $P_{m,n}(F)$ by using approximation (4.3.1) and (appendix)
No
No
Is m even?
Yes
Computing $P_{m,n}(F)$ by using series expansion (4.2.1)
No
Is n even?
Yes
Computing $P_{m,n}(F)$ by using series expansion (4.2.2)
No
Is m = 1?
Yes
Computing $P_{m,n}(F)$ by using series expansion (4.2.5)
No
Is n = 1?
Yes
Computing $P_{m,n}(F)$ by using series expansion (4.2.4)
No
Computing $P_{m,n}(F)$ by using series expansion (4.2.3)
Yes
PUNCH
m, n, F, $P_{m,n}(F)$
STOP

4.5 Test Program

Following is a listing of a Fortran IV program which computes the probability when m, n F and the tabular values are read in. A difference in probability values is computed in order to determine whether or not computation by the calculation procedures presented in the previous section is worthwhile.

The output for the program gives probability obtained by machine computation and the difference (comparing the computed answer P with known answer tabular P). The output shows that absolute error is between 0.00000 and 0.00093 and hence the composite procedure is quite accurate.

Fortran program for the cumulative
F distribution

```
C       THIS IS A TEST PROGRAM TO CHECK THE ANSWERS OBTAINED BY
C       MACHINE COMPUTATION FOR NUMERICAL APPROXIMATIONS WITH
C       THE KNOWN ANSWERS IN THE TABLE
C       BY GRACE YUAN-CHUEN WANG, UTAH STATE UNIVERSITY
        WRITE(3,10)
   10   FORMAT(1H1,7X,6HOUTPUT// )
        WRITE(3,11)
   11   FORMAT (5X,2HV1,4X,2HV2,6X,1HF,9X,5HEST P,5X,5HTAB P,4X,
       14HDIFF//)
  100   READ(2,20) V1,V2,F,TRUEP
   20   FORMAT(2F5.0,2F10.5)
        CALL PROBF(V1,V2,F,ESTP)
        ERROR=TRUEP-ESTP
        JV1=V1
        JV2=V2
        WRITE(3,12) JV1,JV2,F,ESTP,TRUEP,ERROR
   12   FORMAT(2X,I5,1X,I5,2X,4F10.5)
        GO TO 100
        END
```

```
      SUBROUTINE PROBF (V1,V2,F,P)
C     THE CUMULATIVE F DISTRIBUTION
C     COMPUTING PROBABILITY WHEN V1,V2, and F ARE READ IN.
C     V1 IS THE NUMBER OF DEGREES OF FREEDOM OF THE
C     NUMERATOR OF F
C     V2 IS THE NUMBER OF DEGREES OF FREEDOM OF THE
C     DENOMINATOR OF F
      JV1=V1
      JV2=V2
      IF(F-1.01) 101,101,130
  130 IF(JV1-8) 101,102,102
  102 IF(JV2-8) 101,103,103
C     COMPUTING P WHEN BOTH V1 AND V2 ARE GREAT THAN 8 OR
C     EQUAL TO 8
  103 A=2./(9.*V2)
      B=2./(9.*V1)
      X=(F**(1./3/)*(1.-A)-(1.-B) )/SQRT  (B+F**(2./3.)*A)
      ZX=1./(2.5052367*1.2831204*X*X+.2264718*X**4+.1306469*X**6
     1-.0202490*X**8+.0039132*X**10)
      C=1./(1.+.2316419*X)
      P=1.-ZX*(.319381530*C-.356563782*C*C+1.781477937*C**3
     1-1.821255978*C**4+1.330274429*C**5)
      GO TO 99
  101 IF( (-1)**JV1)106,106,105
C     COMPUTING P WHEN V1 IS EVEN AND LESS THAN 8
  105 X=V2/(V2+V1*F)
      IF(JV1-2)27,27,28
   27 P=1.-X**(V2/2.)
      GO TO 99
   28 S=1.0
      SUM=0.0
      K=(JV1-2)/2
      DO 20  N=1,K
      PN=N
      S=S*(V2+2.*PN-2.)/(2.*PN)
   20 SUM=SUM+S*(1.-X)**N
      P=1.-X**(V2/2.)*(1.+SUM)
      GO TO 99
  106 IF( (-1)**JV2)108,108,109
C     COMPUTING P WHEN V2 IS EVEN AND LESS THAN 8
  109 X=V2/(V2+V1*F)
      IF(JV2-2)21,21,22
   21 P=(1.-X)**(V1/2.)
      GO TO 99
   22 R=1.
      TOTAL=0.
      M=(JV2-2)/2
      DO 24 I=1,M
```

```
          RI=I
          R=R*(V1+2.*RI-2.)/(2.*RI)
   24     TOTAL=TOTAL+R*X**I
          P=(1.-X)**(V1/2.)*(1.+TOTAL)
          GO TO 99
      108 IF(JV1-1) 120,120,121
C         COMPUTING P WHEN V1=1 AND V2 IS ODD NUMBER
      120 T=SQRT (F)
          B=ATAN (T/SQRT (V2) )
          IF(V2-3.)35,36,37
       35 P=2.*B/3.141593
          GO TO 99
       36 P=2./3.141593*(B+SIN (B)*COS  (B) )
          GO TO 99
       37 A=1.
          SUM=0.
          J=JV2-2
          DO 30 N=3,J,2
          PN=N
          A=A*(PN-1.)/PN
       30 SUM=SUM+A*COS (B)**N
          P=2./3.141593*(B+SIN (B)*COS (B)+SIN (B)*SUM)
          GO TO 99
      121 IF(JV2-1) 122,122,123
C         COMPUTING P WHEN V2=1 and V1 IS ODD NUMBER EXCEPT V1=1
      122 Y=SQRT (V1/V2*F)
          X=ATAN (Y)
          D=2./3.141593*SIN (X)*COS (X)
          IF (JV1-3) 48,48,49
       48 BV1V2=D
          GO TO 44
       49  Z=1.
          SUMW=0.
          K=JV1-3
          DO 40 I=2,K,2
          PI=I
          Z=Z*(V2+PI-1.)/(PI+1.)
       40 SUMW=SUMW+Z*SIN (X)**I
          BV1V2=D*(1.+SUMW)
       44 P=2./3.141593*X-BV1V2
          GO TO 99
C         COMPUTING P WHEN BOTH V1 AND V2 ARE ODD, V1 OR V2 IS LESS
C         THAN 8 AND BOTH V1 AND V2 ARE NOT EQUAL TO 1
      123 Y=SQRT (V1/V2*F)
          X=ATAN (Y)
          IF(JV2-3) 56,56,57
       56 ATV2=2./3.1416*(X+SIN (X)*COS (X) )
          GO TO 77
```

```
57 A=1.
   SUM=0.
   J=JV2-2
   DO 50 N=3,J,2
   PN=N
   A=A*(PN-1.)/PN
50 SUM=SUM+A*COS (X)**N
   ATV2=2./3.1416*(X+SIN (X)*COS (X)+SIN (X)*SUM)
77 K=(JV2-1)/2
   NN=1
   DO 52 N=1,K
52 NN=NN*N
   UPFAC=NN
   L=(JV2-3)/2
   IF(L) 51,51,54
51 BUFA=1./2.*1.77245
   GO TO 55
54 M=2*L+1
   KK=1
   DO 61 I=1,M,2
61 KK=KK*I
   R=KK
   E=L+1
   BUFA=(R/2.**E)*1.77245
55 T=2./1.77245*UPFAC/BUFA
   IF(JV1-3) 58,58,59
58 BV1V2=T*SIN (X)*COS (X)**JV2
   GO TO 66
59 W=1.
   SUMW=0.
   K=JV1-3
   DO 70 I=2,K,2
   PI=I
   W=W*(V2+PI-1.)/(PI+1.)
70  SUMW=SUMW+W*SIN (X)**I
   BV1V2=T*SIN (X)*COS (X)**JV2*(1.+SUMW)
66 P=ATV2-BV1V2
99 RETURN
   END
```

OUTPUT

V1	V2	F	EST P	TAB P	DIFF
1	3	10.10000	.94983	.95000	.00017
1	7	5.59000	.94998	.95000	.00002
5	1	230.00000	.94998	.95000	.00002
3	3	9.28000	.95002	.95000	-.00002
5	3	9.01000	.94997	.95000	.00003
5	7	3.97000	.94995	.95000	.00005
2	5	5.79000	.95006	.95000	-.00006
2	6	5.14000	.94994	.95000	.00006
6	6	4.28000	.94990	.95000	.00010
1	2	18.50000	.94997	.95000	.00003
3	6	4.76000	.95006	.95000	-.00006
7	4	6.09000	.94994	.95000	.00006
5	6	4.39000	.95006	.95000	-.00006
6	16	2.74000	.94992	.95000	.00008
3	20	3.10000	.95008	.95000	-.00008
60	6	3.74000	.95001	.95000	-.00001
12	8	3.28000	.94941	.95000	.00059
9	10	3.02000	.94980	.95000	.00020
40	24	1.89000	.94973	.95000	.00027
40	30	1.79000	.94973	.95000	.00027
20	60	1.75000	.95044	.95000	-.00044
12	120	1.83000	.94963	.95000	.00037
7	3	8.89000	.95002	.95000	-.00002
7	7	3.79000	.95009	.95000	-.00009
2	7	4.74000	.95005	.95000	-.00005
4	4	6.39000	.95002	.95000	-.00002
6	1	234.00000	.95000	.95000	-.00000
6	4	6.16000	.94996	.95000	.00004
4	1	225.00000	.95005	.95000	-.00005
5	2	19.30000	.95001	.95000	-.00001
1	4	7.71000	.95001	.95000	-.00001
7	6	4.21000	.95009	.95000	-.00009
120	1	253.00000	.94998	.95000	.00003
12	2	19.40000	.94997	.95000	.00003
40	3	8.59000	.94996	.95000	.00004
24	5	4.53000	.95006	.95000	-.00006
20	6	3.87000	.94987	.95000	.00013
1	11	4.84000	.94991	.95000	.00009
6	30	2.42000	.94996	.95000	.00004
2	120	3.07000	.94992	.95000	.00008
24	18	2.15000	.95000	.95000	.00000
60	18	2.02000	.95032	.95000	-.00032

15	22	2.15000	.94996	.95000	.00004
120	22	1.84000	.95026	.95000	-.00026
5	1	.10000	.02503	.02500	-.00003
2	5	.30500	.25008	.25000	-.00008
30	8	.53100	.10016	.10000	-.00016
9	9	4.03000	.97449	.97500	.00051
8	15	.61800	.25968	.25000	.00032
1	3	.00120	.02546	.02500	-.00046
1	3	.12200	.25005	.25000	-.00005
1	3	2.02000	.74964	.75000	.00036
1	3	167.00000	.99900	.99900	0.00000
1	9	.00017	.01012	.01000	-.00012
1	9	.01700	.10087	.10000	-.00087
1	9	.10800	.25005	.25000	-.00005
1	15	.00100	.02481	.02500	.00019
1	15	.10500	.24962	.25000	.00038
1	15	10.80000	.99500	.99500	-.00000
3	1	.05700	.02481	.02500	.00019
7	1	.02700	.00050	.00050	.00000
11	1	.05100	.00101	.00100	-.00001
11	1	9.36000	.74991	.75000	.00009
4	4	.04300	.00496	.00500	.00004
4	4	1.00000	.50000	.50000	0.00000
4	4	53.40000	.99900	.99900	.00000
5	5	.02500	.00049	.00050	.00001
5	5	14.90000	.99497	.99500	.00003
8	8	.08300	.00100	.00100	.00000
8	8	1.00000	.50000	.50000	0.00000
12	12	4.16000	.98975	.99000	.00025
30	4	.70200	.24998	.25000	.00002
20	2	.38600	.09989	.10000	.00011
6	20	6.02000	.99900	.99900	-.00000
60	11	.51300	.05038	.05000	-.00038
120	7	.32400	.00501	.00500	-.00001
7	30	.06500	.00051	.00050	-.00001
2	120	5.54000	.99500	.99500	-.00000
200	60	.77600	.10028	.10000	-.00028
500	24	3.01000	.99890	.99900	.00010
200	40	.48000	.00051	.00050	-.00001
9	8	1.01000	.50015	.50000	-.00015
10	120	.93900	.49962	.50000	.00038
100	8	1.08000	.49906	.50000	.00094
0	0	0.00000	0.00000	0.00000	0.00000

REFERENCES

Elderton, W. P. 1902. Biometrika I. Cambridge University Press, New York. p. 155-163.

Fisher, R. H. 1925. Statistical Methods for Research Worker. Oliver and Boyd, Edenburgh.

Hastings, C. Jr., and J. Wong, Jr. 1954. Mathematical Tables and Other Aids to Computation. National Research Council. January.

Hastings, Cecil, Jr. 1955. Approximation for Digital Computers. Princeton University Press, Princeton, New Jersey.

Wilson, E. B., and M. M. Hilferty. 1931. Proceedings of the National Academy of Science, Washington. 17:684

Zelen, Marvin, and Norman C. Serevo. 1965. Probability Function, Handbook of Mathematical Function. National Bureau of Standards, New York.

APPENDIX

1. Numerical Approximation to the Cumulative Standardized Normal Distribution

The variable u is normally distribution with mean 0 and variance 1. The function

$$Z(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

is called the standardized normal distribution function. The standardized distribution function is symmetrical about $u = 0$ and hence $Z(-u) = Z(u)$.

The cumulative distribution p (u) is equal to

$$p(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{u^2}{2}} du$$

A very accurate approximation to p(u) has been given by Cecil Hastings, Jr. (1955) i. e.,

$$p(u) = 1 - Z(u)\,(b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 + b_5 t^5) \qquad ((A.1.1)$$

$$t = \frac{1}{1 + pu} \qquad p = .2316419$$

$b_1 = 0.319381530$

$b_2 = -0.356563785$

$b_3 = 1.781477939$

$b_4 = -1.821255978$

$b_5 = 1.330274429$

$$Z(u) = (a_0 + a_2u^2 + a_4u^4 + a_6u^6 + a_8u^8 + a_{10}u^{10})^{-1}$$

$a_0 = 2.5052367$

$a_2 = 1.2831204$

$a_4 = 0.2264718$

$a_6 = 0.1306469$

$a_8 = -0.0202490$

$a_{10} = 0.0039132$

2. Numerical Approximation to $\sqrt{\frac{2}{\pi}} \int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2} d\chi$

An accurate approximation to $\frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ has been given by Cecil Hastings, Jr. (1955).

$$\frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt \doteq 1 - (1 - C_1 z^1 + C_2 z^2 + C_3 z^3 + C_4 z^4 + C_5 z^5 + C_6 z^6)^{-16}$$

$C_1 = 0.0706230784$

$C_2 = 0.0422820123$

$C_3 = 0.0092705272$

$C_4 = 0.0001520143$

$C_5 = 0.002765672$

$C_6 = 0.0000430638$

Now we use the simple transformation. Let $\chi = \sqrt{2}t$ and $d\chi = \sqrt{2}\,dt$

Hence $$\sqrt{\frac{2}{\pi}} \int_0^{\chi} e^{-\frac{1}{2}\chi^2} d\chi = \frac{2}{\sqrt{\pi}} \int_0^{\frac{\chi}{\sqrt{2}}} e^{-t^2} dt$$

Since $$\sqrt{\frac{2}{\pi}} \int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2} d\chi = 1 - \sqrt{\frac{2}{\pi}} \int_0^{\chi} e^{-\frac{1}{2}\chi^2} d\chi$$

$$\sqrt{\frac{2}{\pi}} \int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2} d\chi = 1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{\chi}{\sqrt{2}}} e^{-t^2} dt \qquad \text{(A.2.1)}$$

Suppose we need to find the value of $\sqrt{\frac{2}{\pi}} \int_{3}^{\infty} e^{-\frac{1}{2}\chi^2} d\chi$

$$\sqrt{\frac{2}{\pi}} \int_{3}^{\infty} e^{-\frac{1}{2}\chi^2} d\chi = 1 - \frac{2}{\sqrt{\pi}} \int_{0}^{2.12} e^{-t^2} dt \quad \text{(using (A. 2. 2))}$$

$$\frac{2}{\sqrt{\pi}} \int_{0}^{2.12} e^{-t^2} dt = 0.99728 \qquad \text{(using approximation (A. 2. 1))}$$

and hence

$$\sqrt{\frac{2}{\pi}} \int_{3}^{\infty} e^{-\frac{1}{2}\chi^2} d\chi = 0.00272$$