Calculation of critical values of several probability distributions using standard numerical methods

Dušan Knežo[1], Jozef Zajac [2], Peter Michalik[3]

Abstract. Very common statistical methods used in practice are interval estimations, tests of statistical hypotheses and other tests. Using these methods usually requires knowledge of critical values of some distributions. In the case of mentioned application methods are used computer programs, but in other cases it is desirable to have a method of calculating the distribution of the critical values. Article presents methods for calculating the critical values for a standard normal distribution, Student's t-distribution, Fisher-Snedecor F-distribution and χ^2-distribution using standard numerical methods, specifically an approximate calculation of definite integrals and approximate calculation of nonlinear equations.

Keywords: probability distributions, critical values, approximate calculation of definite integrals, approximate calculation of nonlinear equations.

1 Introduction

Application of some statistical methods for example interval estimations or statistical hypothesis tests, which requires some critical values probability distribution, particularly a standard normal distribution, Student's t-distribution, Fisher-Snedecor F-distribution and χ^2-distribution.

If the function $F(x)$ is distribution function of the probability distribution of the random variable X, thus the critical value of this distribution to the significance level α is the value x_α, to which is valid

$$P(X > x_\alpha) = 1 - F(x_\alpha) = \alpha. \qquad (1)$$

Determining critical values is usually made by using tables or specialized statistical software. This approach also has its disadvantages, but in some cases is advantageous to have a method by which is possible to calculate evaluated critical values. Calculation of critical values is also possible using standard numerical methods, specifically method of approximate calculation of definite integrals and methods of approximate calculation of nonlinear equations. Analogical uses of methods are presented in article according [1- 3].

1 Technical University of Košice, Faculty of manufacturing technologies with a seat in Prešov, Department of mathematics, informatics and cybernetics, Bayerova 1, 080 01 Prešov, Slovak Republic , email: dusan.knezo@tuke.sk
2 Technical University of Košice, Faculty of manufacturing technologies with a seat in Prešov, Department of computer aided manufacturing technologies, Štúrova 31, 080 01 Prešov, Slovak Republic , email: jozef.zajac@tuke.sk
3 Technical University of Košice, Faculty of manufacturing technologies with a seat in Prešov, Department of manufacturing technologies, Štúrova 31, 080 01 Prešov, Slovak Republic , email: peter.michalik@tuke.sk

Study [2] prescribes methods for determination critical values of standard normal and Student's t-distribution.

Equation (1) is used to determinate critical values x_α of standard normal distribution N(0,1) in the level of significant α and implies (Fig. 1)

$$1 - \phi(x_\alpha) = \alpha, \qquad (2)$$

where $\phi(x)$ is distribution function of standard normal distribution N(0,1) and is define by equation

$$\phi(x) = \int_{-\infty}^{x} \varphi(t)dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt. \qquad (3)$$

From the properties of the distribution function $\varphi(x)$ implies that

$$\phi(x) = \begin{cases} 0.5 - \frac{1}{\sqrt{2\pi}}\int_0^{|x|} e^{-\frac{t^2}{2}}dt\,, & x < 0, \\ 0.5, & x = 0, \\ 0.5 + \frac{1}{\sqrt{2\pi}}\int_0^{x} e^{-\frac{t^2}{2}}dt\,, & x > 0. \end{cases} \qquad (4)$$

From equation (4) can be state, that is je is sufficient to limit consideration to the integral

$$\int_0^{x} \varphi(t)dt = \frac{1}{\sqrt{2\pi}}\int_0^{x} e^{-\frac{t^2}{2}}\,dt. \qquad (5)$$

valid for $x > 0$. Calculating integral (5) can be used Simpson´s rule and can be deduces

$$\left|\varphi^{(4)}(t)\right| \leq \varphi^{(4)}(0) = \frac{3}{\sqrt{2\pi}}. \qquad (6)$$

For integration step h of approximate calculation of integral (5) using Simpson’s rule with accuracy ε implies

$$h \leq \sqrt[4]{\frac{60\cdot\varepsilon\cdot\sqrt{2\pi}}{x}}. \qquad (7)$$

Based on previous equations is possible with accuracy ε to calculate values of the distribution function, thus is possible to set critical value x_α with required accuracy to evaluate equation

$$\phi(x_\alpha) + \alpha - 1 = 0. \qquad (8)$$

Because function

$$G(x) = \phi(x) + \alpha - 1 \qquad (9)$$

is increasing continuous function,

$$\lim_{x \to -\infty} G(x) = \alpha - 1 < 0 \qquad (10)$$

and

- $\lim_{x \to \infty} G(x) = \alpha > 0, \qquad (11)$

then equation (8) has exactly one solution. Approximate solution of equation (8) can be done using bisection method.
Table 1 shows some critical values of standard normal distributions calculated by described method.
Probability density according to Students *t*-distribution with *n* degrees of freedom is evident, that distribution function can be define as follows

- $F_n(x) = \begin{cases} 0.5 - c_n \cdot \int_0^{|x|} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt, & x < 0, \\ 0.5, & x = 0, \\ 0.5 + c_n \cdot \int_0^{x} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt, & x > 0. \end{cases} \qquad (12)$

where

$$c_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \cdot \Gamma\left(\frac{n}{2}\right)}. \qquad (13)$$

For approximate integral calculation

$$\int_0^x g_n(t)\, dt = \int_0^x \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt, \qquad (14)$$

for $x > 0$ can be used trapezoidal rule. Can be demonstrated

$$|g_n''(t)| \leq g_n''(0) = \frac{n+1}{n}. \qquad (15)$$

For step of approximate integral calculation (14) using trapezoidal rule with accuracy ε is valid

$$h \leq \sqrt{\frac{12n\varepsilon}{(n+1) \cdot x}}. \qquad (16)$$

Determination of critical value x_α by Students *t*-distribution (Fig. 2) is to find the equation root

$$1 - F_n(x) = \alpha. \qquad (17)$$

Approximate solution of equation (17) can be realized using bisection method. Table 2 shows some critical values of Student t-distribution with n degrees of freedom calculated by described method.

2 Critical values of χ^2-distribution

Probability density χ^2-distribution with *n* degrees of freedom is define by equation

$$f_n(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leq 0. \end{cases} \qquad (18)$$

Because for $n \in N$ and $n \geq 2$ for calculation values of function Γ can be used equation

$$\Gamma(k) = (k-1)! \qquad (19)$$

or

$$\Gamma\left(k + \frac{1}{2}\right) = \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{2^k} \cdot \sqrt{\pi}, \qquad (20)$$

where $k \in N$.
Equation (1) for determinate critical value x_α χ^2-distribution (Fig. 3) with *n* degrees of freedom on confidence level α is apparent

$$1 - F_n(x) = \alpha, \qquad (21)$$

where $F_n(x)$ is distribution function define by equation

$$F_n(x) = \int_0^x f_n(t)\, dt. \qquad (22)$$

Derivation properties of function $f_n(x)$ shows, that in general is not possible to define step for approximate integral calculation (22) to obtain required accuracy. In that case is approximate integral calculation (22) would be appropriate to use

Richardson extrapolation and Simpson’s method. Calculation uncertainty/error E is estimate by equation

$$|E| \cong \frac{\left|I_h - I_{h/2}\right|}{15}, \qquad (23)$$

where I_s is approximate value of the integral calculated by the Simson's method with using step s.
Equation (21) has exactly one solution and its approximate values can be found using bisection method. Table 3 shows some critical values for χ^2-distribution with n degrees of freedom calculated using described evaluation method.

3 Critical values of *F*-distribution

The probability density of F-distribution with m, n degrees of freedom is in interval $[0, \infty)$ prescribed by equation

$$f_{m,n}(x) = c_{m,n} \cdot \frac{x^{\frac{m}{2}-1}}{(mx+n)^{\frac{m+n}{2}}}, \qquad (24)$$

where

$$c_{m,n} = \frac{\Gamma\left(\frac{m+n}{2}\right) \cdot m^{\frac{m}{2}} \cdot n^{\frac{n}{2}}}{\Gamma\left(\frac{m}{2}\right) \cdot \Gamma\left(\frac{n}{2}\right)}. \qquad (25)$$

Because $m, n \in N$ values of function Γ can be calculated using equations (19) and (20).

Equation (1) for determination of critical values x_α, subsequently F-distribution with m, n degrees of freedom on significant level α is apparent (Fig. 4).

$$1 - F_{m,n}(x_\alpha) = \alpha, \qquad (26)$$

where $F_{m,n}(x)$ is distribution function and is define by equation

$$F_{m,n}(x) = \int_0^x f_{m,n}(t)\, dt. \qquad (27)$$

Analogy with the previous section, in general is not possible to determined step of approximate integral calculation (27) to obtain a specified accuracy. In this case it is also appropriate to use Richardson extrapolation method and Simpson’s method.

Equation (26) has exactly one solution and approximate value can be possible find using bisection method. Table 4-7 list some of critical values of F-distribution with m, n degrees of freedom calculated using prescribed method.

4 Conclusions

Described methods present possibilities of calculation critical values for the most common probability distributions used in practice. Advantage of these methods is that they can be used as part of automatic data processing, e.g. at interval estimates or statistical testing. Advantageous is that they allows the calculation of the critical values for any significance level. Methods can also be used to create tables included in books and textbooks in the field of statistics. Program for each of described methods was designed in a language Free Pascal - Lazarus IDE. Algorithms of these programs are of course usable in other programming languages. Application of described methods is limited with technical possibilities of computer, because at some values computer works at machine's boundary limits. Described methods are based on standard numerical methods of calculation definite integrals. Other methods such as Monte Carlo and others can be used as well.

Acknowledgements

This work is a part of research project VEGA 1/0619/15. The article was prepared within the invitation OPVaV-2012/2.2/08-RO named "University Science Park TECHNICOM for Innovation Applications with the Support of Knowledge-Based Technologies" code ITMS 26220220182 and of the project VEGA 1/0594/12.

References

[1] KNEŽO, D. (2014) Inverse transformation method for normal distribution and the standard numerical method. International Journal of Interdisciplinarity in Theory and Practice 5: 6-10.

[2] KNEŽO, D. (2014) Application of the numerical methods for the teaching of the statistics. International Journal of Interdisciplinarity in Theory and Practice 5: 58-61.

[3] KNEŽO, D. (2012) O metóde Monte Carlo a možnostiach jej aplikácií. Transfer inovácií 24: 178-181.

[4] BONNAR J. (2014) J. Bonnar, The Gamma Function (CreateSpace).

[5] WALCK CH. (2007) Hand-book on statistical distributions for experimentalist (Universitet Stockholms).

[6] KOKOSKA S., NEVISON CH. (1992) Statistical Tables and Formulae (New York: Springer-Verlag).
[7] NEAVE H. R. (1981) Elementary Statistics Tables (London: George Allen & Unwin).
[8] NEAVE H. R. (2011) Elementary Statistics Table (London: Routledge).
[9] NEAVE H. R. (2011) Statistics Tables: For Mathematicians, Engineers, Economists and the Behavioural and Management Sciences (London: Routledge)

Figure 1. Critical values of standard normal distribution [6]

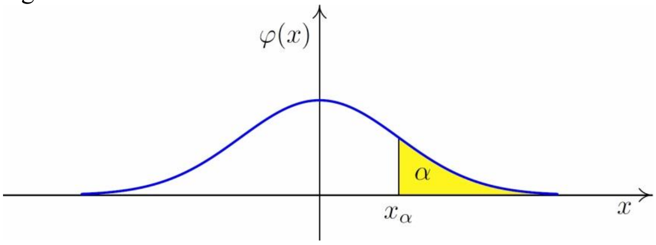

Figure 2. Critical values of Students *t*-distribution [7]

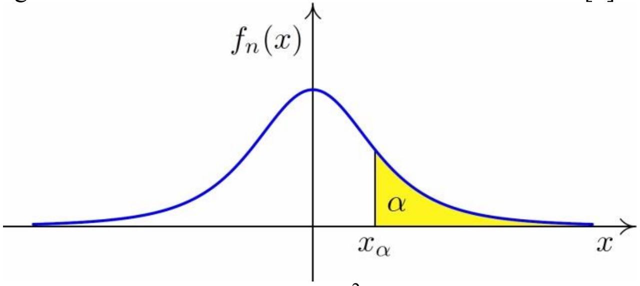

Figure 3. Critical values of χ^2-distribution

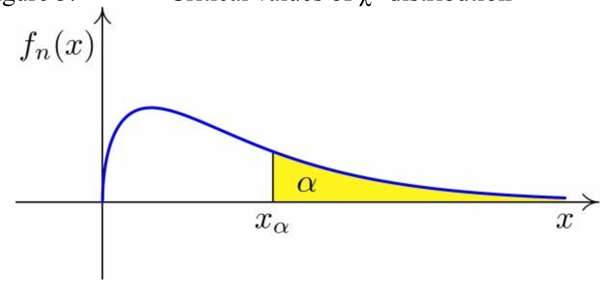

Figure 4. Critical values of F-distribution

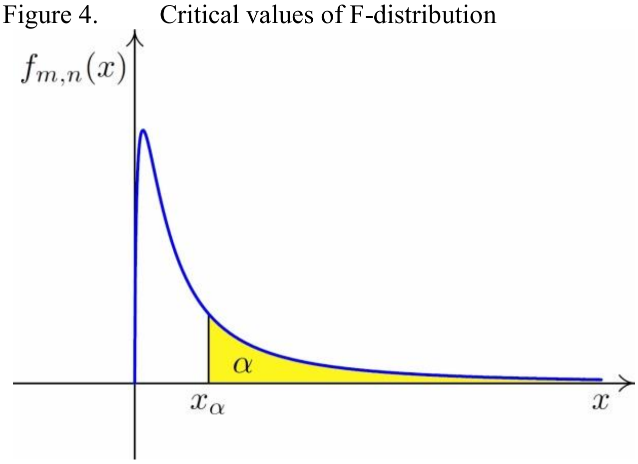

Table 1. Selected critical values of standard normal distribution

α	0.10	0.05	0.025	0.0125	0.01	0.005
x_α	1.281551	1.644854	1.959965	2.241401	2.326350	2.575833

Table 2. Selected critical values of Students t-distribution

n	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.001$
3	1.638	2.353	3.182	4.541	10.215
4	1.533	2.132	2.776	3.747	7.173
5	1.476	2.015	2.571	3.365	5.893
6	1.440	1.943	2.447	3.143	5.208
7	1.415	1.895	2.365	2.998	4.785
8	1.397	1.860	2.306	2.896	4.501
9	1.383	1.833	2.262	2.821	4.297
10	1.372	1.812	2.228	2.764	4.144
11	1.363	1.796	2.201	2.718	4.025
12	1.356	1.782	2.179	2.681	3.930
13	1.350	1.771	2.160	2.650	3.852
14	1.345	1.761	2.145	2.624	3.787
15	1.341	1.753	2.131	2.602	3.733
16	1.337	1.746	2.120	2.583	3.686
17	1.333	1.740	2.110	2.567	3.646
18	1.330	1.734	2.101	2.552	3.611
19	1.328	1.729	2.093	2.539	3.579
20	1.325	1.725	2.086	2.528	3.552
21	1.323	1.721	2.080	2.518	3.527
22	1.321	1.717	2.074	2.508	3.505
23	1.319	1.714	2.069	2.500	3.485

Table 3. Selected critical values of χ^2-distribution

n	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.001$
3	6.251	7.815	9.348	11.345	16.267
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.070	12.832	15.086	20.515
6	10.645	12.592	14.449	16.812	22.458
7	12.017	14.067	16.013	18.475	24.322

8	13.362	15.507	17.535	20.090	26.125
9	14.684	16.919	19.023	21.666	27.877
10	15.987	18.307	20.483	23.209	29.588
11	17.275	19.675	21.920	24.725	31.264
12	18.549	21.026	23.337	26.217	32.910
13	19.812	22.362	24.736	27.688	34.528
14	21.064	23.685	26.119	29.141	36.123
15	22.307	24.996	27.488	30.578	37.697
16	23.542	26.296	28.845	32.000	39.252
17	24.769	27.587	30.191	33.409	40.791
18	25.989	28.869	31.526	34.805	42.312
19	27.204	30.144	32.852	36.191	43.820
20	28.412	31.410	34.170	37.566	45.315
21	29.615	32.671	35.479	38.932	46.797
22	30.813	33.924	36.781	40.289	48.268
23	32.007	35.172	38.076	41.638	49.729

Table 4. Selected critical values of F-distribution for $\alpha = 0.1$

	m=3	*m*=4	*m*=5	*m*=6	*m*=7	*m*=8	*m*=9	*m*=10
n=3	5.391	5.343	5.309	5.285	5.266	5.253	5.240	5.230
n=4	4.191	4.107	4.051	4.010	3.979	3.955	3.936	3.920
n=5	3.620	3.520	3.453	3.404	3.368	3.339	3.316	3.297
n=6	3.289	3.181	3.107	3.055	3.014	2.983	2.958	2.937
n=7	3.074	2.961	2.883	2.827	2.785	2.752	2.725	2.702
n=8	2.924	2.806	2.726	2.668	2.624	2.589	2.561	2.538
n=9	2.813	2.693	2.611	2.551	2.505	2.469	2.440	2.416
n=10	2.728	2.605	2.522	2.461	2.414	2.377	2.347	2.323

Table 5. Selected critical values of F-distribution for $\alpha = 0.05$

	m=3	*m*=4	*m*=5	*m*=6	*m*=7	*m*=8	*m*=9	*m*=10
n=3	9.278	9.117	9.013	8.941	8.887	8.845	8.812	8.786
n=4	6.592	6.388	6.256	6.163	6.094	6.041	5.999	5.964
n=5	5.410	5.192	5.050	4.950	4.876	4.818	4.779	4.735
n=6	4.757	4.534	4.387	4.284	4.207	4.148	4.099	4.060
n=7	4.347	4.120	3.971	3.866	3.787	3.726	3.677	3.637
n=8	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
n=9	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
n=10	3.708	3.478	3.326	3.217	3.136	3.072	3.020	2.978

Table 6. Selected critical values of F-distribution for $\alpha = 0.025$

	m=3	m=4	m=5	m=6	m=7	m=8	m=9	m=10
n=3	15.442	15.101	14.884	14.734	14.625	14.540	14.473	14.419
n=4	9.980	9.605	9.364	9.197	9.074	8.980	8.905	8.844
n=5	7.764	7.388	7.146	6.978	6.854	6.757	6.681	6.619
n=6	6.599	6.227	5.987	5.820	5.696	5.600	5.523	5.461
n=7	5.890	5.523	5.285	5.119	4.996	4.899	4.823	4.761
n=8	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295
n=9	5.079	4.718	4.484	4.320	4.197	4.102	4.026	3.964
n=10	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717

Table 7. Selected critical values of F-distribution for $\alpha = 0.01$

	m=3	m=4	m=5	m=6	m=7	m=8	m=9	m=10
n=3	29.468	28.710	28.235	27.910	27.682	27.490	27.346	27.229
n=4	16.700	15.977	15.521	15.123	14.980	14.799	14.659	14.546
n=5	12.062	11.392	10.966	10.672	10.456	10.290	10.158	10.051
n=6	9.783	9.148	8.745	8.466	8.260	8.102	7.976	7.874
n=7	8.453	7.847	7.460	7.191	6.993	6.840	6.719	6.620
n=8	7.592	7.006	6.632	6.371	6.178	6.029	5.911	5.814
n=9	6.993	6.422	6.056	5.802	5.615	5.467	5.351	5.257
n=10	6.553	5.994	5.636	5.386	5.200	5.057	4.942	4.849