

Figure 3: Chi square distributions with different degrees of freedom

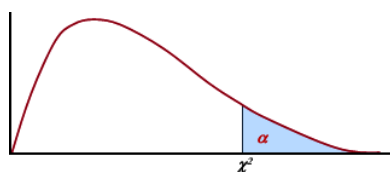


Figure 4: χ^2 distribution with ν degrees of freedom

2 Lecture 2

2.1 The chi square distribution

In particular, when $\alpha = \nu/2$ and $\beta = 2$, we have the **chi square distribution** (χ^2) with ν degrees of freedom.

We use $\chi^2(\nu)$ to denote a random variable having a chi square distribution with ν degrees of freedom, Figure 3. Note that random variables $\chi^2(\nu)$ can only take on non-negative values. So their probability density function is **not** symmetric.

Example 7 Show that the probability density function of a chi square distribution with ν degrees of freedom is

$$f(x) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{\frac{\nu}{2}-1}e^{-x/2}, & x > 0 \\ 0 & x \leq 0, \end{cases} \quad (23)$$

where the gamma function Γ is given by Eq. (18).

Solution. Eq. (23) is an immediate consequence of Eq. (17).

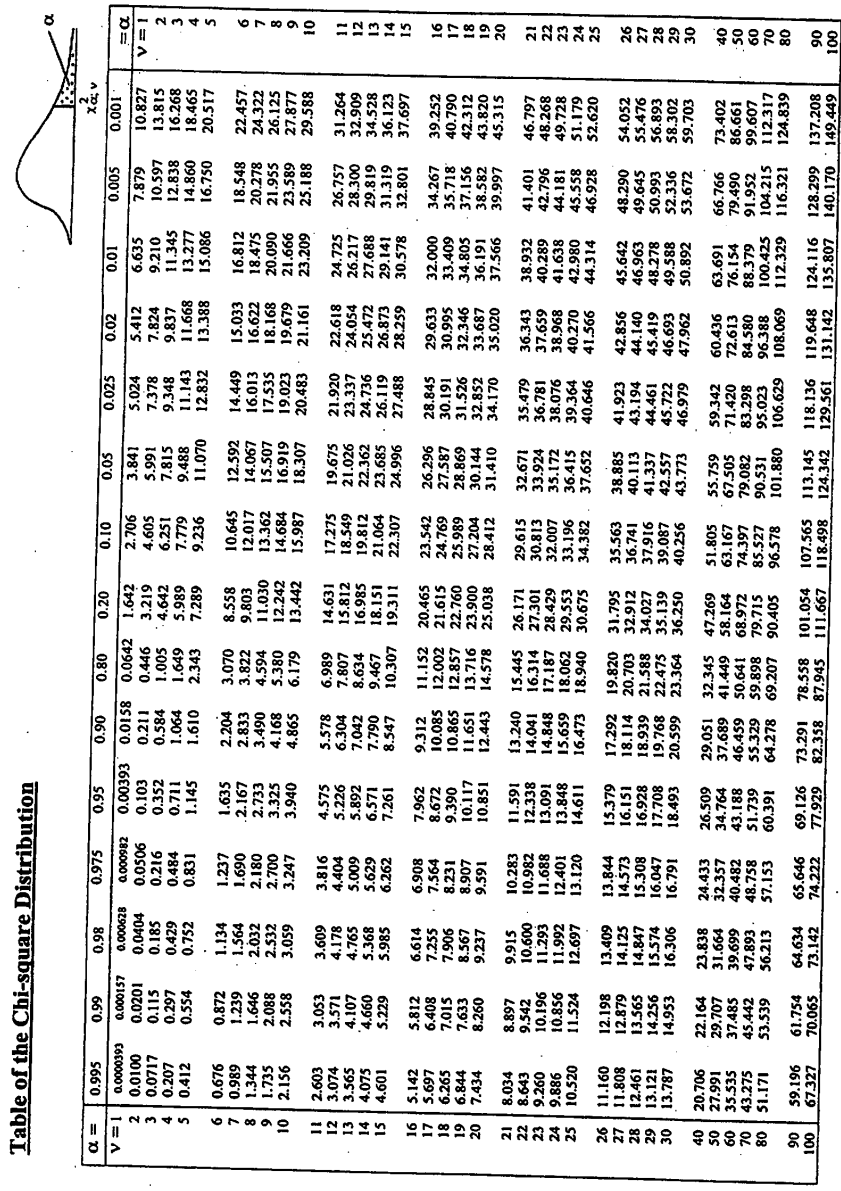


Figure 5: Chi square distribution table

Let χ_α^2 be such that the area under the chi square distribution to its right is equal to α , see the tail in Figure 4. That is

$$P(\chi^2(\nu) \geq \chi_\alpha^2) = \alpha. \quad (24)$$

Example 8 Let a random variable χ^2 have a chi square distribution with 19 degree of freedom, find $\chi_{0.05}^2$ and $\chi_{0.01}^2$.

Solution. $\nu = 19$. By the chi square table (Figure 5),

$$\chi_{0.05}^2 = 30.14, \quad \chi_{0.01}^2 = 36.19.$$

Example 9 Let $\alpha = 0.05$. Evaluate $P(\chi_{1-\alpha/2}^2 < \chi^2(\nu) < \chi_{\alpha/2}^2)$.

Solution. Since

$$P(\chi^2(\nu) \geq \chi_{\alpha/2}^2) = \frac{\alpha}{2}, \quad P(\chi^2(\nu) \geq \chi_{1-\alpha/2}^2) = 1 - \frac{\alpha}{2},$$

$$P(\chi_{1-\alpha/2}^2 < \chi^2(\nu) < \chi_{\alpha/2}^2) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - 0.05 = 0.95.$$

(draw a plot for the example)

2.2 Representation of Chi Square Random Variable

Theorem 3 Let Z_1, Z_2, \dots, Z_ν be independent standard normal random variables, where ν is a positive integer. Then the random variable

$$\chi^2(\nu) = \sum_{i=1}^{\nu} Z_i^2 \quad (25)$$

has a chi square distribution with ν degrees of freedom.

The sum of two independent chi square variables, $\chi^2(\nu_1) + \chi^2(\nu_2)$, has chi square distribution with degrees of freedom of $\nu_1 + \nu_2$.

Example 10 If Z has a standard normal distribution, find the probability $P(Z^2 > 7.879)$.

Solution. By Theorem 3, Z^2 has a chi square distribution with one degree of freedom.

Therefore, by the chi square distribution table, we find $P(Z^2 > 7.879) = 0.005$.

Example 11 Let X_1, X_2, \dots, X_n be independent normal random variables all having the same mean μ and variance σ^2 . Define

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (26)$$

and

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

Show that

- (a) Z has a standard normal distribution.
- (b) Z^2 has a chi square distribution with one degree of freedom.
- (c) Find $P(Z^2 > 5.024)$. Does this value depend on n, μ , or σ^2 ? Why?

Solution.

- (a). By Theorem 2, Z is a standard normal random variable.
- (c). by Theorem 3, Z^2 has a chi square distribution with one degree of freedom.
- (c). $\nu = 1$. By the chi square distribution table, $P(Z^2 > 5.024) = 0.025$.

Example 12 Let Z_1, Z_2, \dots, Z_n be independent standard normal distribution. Show that

(a)

$$\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n (Z_i - \bar{Z})^2 + n\bar{Z}^2. \quad (27)$$

- (b) $\sqrt{n}\bar{Z}$ is a standard normal random variable;
- (c) $n\bar{Z}^2$ has a chi square distribution with one degree of freedom;

(d) assume that the two terms on the right-hand side of (27) are independent, so $\sum_{i=1}^n (Z_i - \bar{Z})^2$ has a chi square distribution with $n - 1$ degrees of freedom.

Solution.

(a)

$$\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n (Z_i - \bar{Z} + \bar{Z})^2 = \sum_{i=1}^n (Z_i - \bar{Z})^2 + 2 \sum_{i=1}^n (Z_i - \bar{Z})\bar{Z} + n\bar{Z}^2 = \sum_{i=1}^n (Z_i - \bar{Z})^2 + n\bar{Z}^2.$$

(b). According to Theorem 2, \bar{Z} is a normal random variable with mean 0 and variance $\frac{1}{n}$. Standardization gives

$$\sqrt{n}\bar{Z} = \frac{\bar{Z} - 0}{\sqrt{1/n}}.$$

Hence $\sqrt{n}\bar{Z}$ is a standard normal random variable.

(c). By Theorem 3, $n\bar{Z}^2$ has a chi square distribution with one degree of freedom.

(d). According to Theorem 3, $\sum_{i=1}^n Z_i^2$ has a chi square distribution with n degree of freedom. From part (c) above we have also known that $n\bar{Z}^2$ has a chi square distribution with one degree of freedom. Therefore, $\sum_{i=1}^n (Z_i - \bar{Z})^2$ has a chi square distribution with $n - 1$ degrees of freedom.

Finally, if we define

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^2, \quad (28)$$

we conclude from the previous example that $(n-1)S^2$ has a chi square distribution with $n-1$ degrees of freedom. That is, the scaling $\frac{1}{n-1}$ does not change the nature of the distribution, although it changes the specific form of the probability density function.

The above discussions can be generalized to the following result.

Theorem 4 Let X_1, X_2, \dots, X_n be independent normal random variables all having the same mean μ and variance σ^2 . Then

$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (29)$$

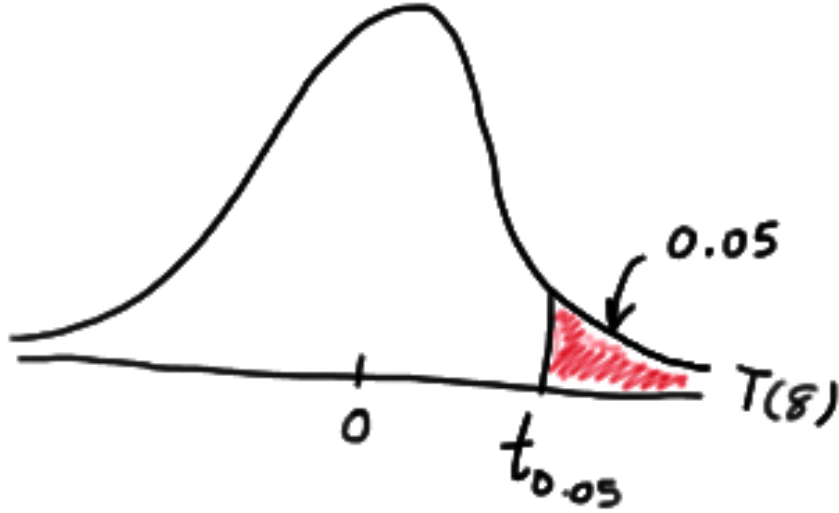


Figure 6: t-distribution with ν degrees of freedom, bell-shaped and symmetric has a χ^2 distribution with $n - 1$ degrees of freedom. Also, \bar{X} and S^2 are independent random variables.

Interpretation of Theorem 4. Note that

$$\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} - \frac{\bar{X} - \mu}{\sigma} \right)^2.$$

Moreover,

$$\frac{X_i - \mu}{\sigma} \sim N(0, 1), \quad \frac{\bar{X} - \mu}{\sigma} \sim N\left(0, \frac{1}{n}\right).$$

Consequently, $\frac{X_i - \mu}{\sigma}$ is like Z_i and $\frac{\bar{X} - \mu}{\sigma}$ is like \bar{Z} in Example 12. So, by Eq. (28), $\frac{(n-1)S^2}{\sigma^2}$ has a χ^2 distribution with $n - 1$ degrees of freedom.

2.3 Representation of t random variable

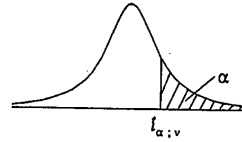
Definition 1 Let the standard normal variable Z and chi square variable χ^2 defined in Eq. (25) be independent. Then

$$t = \frac{\text{standard normal}}{\sqrt{\frac{\text{chi square}}{\text{degrees of freedom}}}} = \frac{Z}{\sqrt{\frac{\chi^2(\nu)}{\nu}}} = \frac{Z}{\sqrt{\frac{\sum_{i=1}^{\nu} Z_i^2}{\nu}}} \quad (30)$$

is said to have a t-distribution with ν degrees of freedom.

Table of the Student's *t*-distribution

The table gives the values of $t_{\alpha;v}$ where
 $\Pr(T_v > t_{\alpha;v}) = \alpha$, with v degrees of freedom



$v \backslash \alpha$	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Figure 7: *t*-distribution

We use t to denote a random variable having a t-distribution.

Let t_α be such that the area under the t-distribution curve to its right is equal to α , see the tail in Figure 6. That is

$$P(t \geq t_\alpha) = \alpha.$$

Example 13 *Let a random variable t have a t distribution with 19 degree of freedom, find $t_{0.05}$.*

Solution. $\nu = 19$. By the t-distribution table (Figure 7),

$$t_{0.05} = 1.729.$$

Example 14 *Let a random variable t have a t distribution with 19 degree of freedom, find $F(-1.729)$.*

Solution. Because t-distribution is **symmetric** (Figure 6),

$$F(-1.729) = P(t \leq -1.729) = P(t \geq 1.729).$$

By the previous example, for $\nu = 19$, $P(t \geq 1.729) = 0.05$. Therefore, $F(-1.729) = 0.05$.

Example 15 *Let Z_1, \dots, Z_6 be independent each having a standard normal distribution. What is the distribution of the random variable*

$$\frac{Z_1 - Z_2}{\sqrt{\frac{Z_3^2 + Z_4^2 + Z_5^2 + Z_6^2}{2}}}$$

Solution. Firstly, since Z_1 and Z_2 are independent standard normal random variables, by the property

$$Var(a_1X_1 + a_2X_2) = a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2, \tag{31}$$

(see Eq. (11)) we find that the random variable

$$\frac{Z_1 - Z_2}{\sqrt{2}}$$

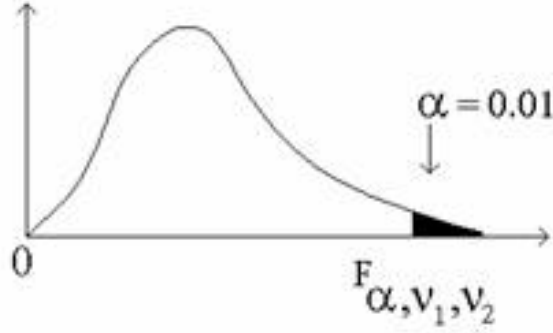


Figure 8: F distribution with ν_1 and ν_2 degrees of freedom

is a standard norm variable. Secondly, note that

$$\frac{Z_1 - Z_2}{\sqrt{\frac{Z_3^2 + Z_4^2 + Z_5^2 + Z_6^2}{2}}} = \frac{\frac{Z_1 - Z_2}{\sqrt{2}}}{\sqrt{\frac{Z_3^2 + Z_4^2 + Z_5^2 + Z_6^2}{4}}}$$

Therefore by Eq. (30), $\frac{Z_1 - Z_2}{\sqrt{\frac{Z_3^2 + Z_4^2 + Z_5^2 + Z_6^2}{2}}}$ has t distribution with 4 degrees of freedom.

2.4 Representation of F random variables

Definition 2 Let the chi square variables $\chi^2(\nu_1)$, with ν_1 degrees of freedom, and $\chi^2(\nu_2)$, with ν_2 degrees of freedom, be independent. Then

$$F(\nu_1, \nu_2) = \frac{\frac{\text{chi square}}{\text{degrees of freedom}}}{\frac{\text{chi square}}{\text{degrees of freedom}}} = \frac{\frac{\chi^2(\nu_1)}{\nu_1}}{\frac{\chi^2(\nu_2)}{\nu_2}} = \frac{\frac{\sum_{i=1}^{\nu_1} Z_i^2}{\nu_1}}{\frac{\sum_{i=\nu_1+1}^{\nu_1+\nu_2} Z_i^2}{\nu_2}} \quad (32)$$

is said to have an F distribution with (ν_1, ν_2) degrees of freedom.

We use $F(\nu_1, \nu_2)$ to denote a random variable having an F -distribution with (ν_1, ν_2) degrees of freedom .

An important identity. Given $0 \leq \alpha \leq 1$, Let $F_\alpha(\nu_1, \nu_2)$ be such that the area under the F -distribution (with degrees of freedom (ν_1, ν_2)) curve to its right is equal to α , see the tail in Figure 8. Then

$$F_{1-\alpha}(\nu_1, \nu_2) = \frac{1}{F_\alpha(\nu_2, \nu_1)}. \quad (33)$$

Example 16 Find the value of $F_{0.95}(10, 20)$.

Table 6(a) Values of F_{α}

$\nu_1 =$ Degrees of Freedom for Denominator	$\nu_2 =$ Degrees of Freedom for Numerator																			
	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120	∞	
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254	
2	1851	1900	1916	1925	1930	1933	1935	1937	1938	1940	1941	1943	1945	1946	1947	1948	1949	1950	1950	
3	1013	955	928	912	901	894	889	885	881	879	878	876	875	874	873	872	871	870	869	
4	7271	694	659	639	626	616	609	604	600	596	591	586	580	577	575	572	569	566	563	
5	6661	579	541	519	505	495	488	482	477	474	468	462	456	452	450	446	443	440	437	
6	599	514	476	453	439	428	421	415	410	406	400	394	387	383	381	377	374	370	367	
7	539	474	438	412	397	384	370	357	343	329	315	301	287	273	259	245	231	217	203	
8	532	466	407	384	369	358	350	342	334	326	318	310	302	294	286	278	270	262	254	
9	512	426	386	363	348	337	329	321	313	305	297	289	281	273	265	257	249	241	233	
10	496	410	371	348	333	322	314	307	302	298	291	285	277	271	266	262	258	254	250	
11	484	398	359	336	320	309	301	295	290	285	279	272	265	260	257	253	249	245	240	
12	475	389	349	326	311	300	291	285	280	275	269	262	254	250	247	243	238	234	229	
13	467	381	341	318	303	292	283	277	271	267	260	253	246	241	238	234	230	226	221	
14	460	374	334	311	296	285	276	270	265	260	253	246	239	234	231	227	222	218	213	
15	454	368	329	306	290	279	271	264	259	254	248	240	233	228	225	220	216	211	207	
16	449	363	324	301	285	274	266	259	254	249	242	235	228	223	219	215	211	206	201	
17	445	359	320	296	281	270	261	255	249	245	238	231	223	218	215	210	206	201	196	
18	441	355	316	293	277	266	258	251	246	241	234	227	219	214	211	206	202	197	192	
19	438	352	313	290	274	263	254	248	242	238	231	223	216	211	207	203	198	193	188	
20	435	349	310	287	271	260	251	245	239	235	228	220	212	207	204	199	195	190	184	
21	432	347	307	284	268	257	249	242	237	232	225	218	210	205	201	196	192	187	181	
22	430	344	305	282	266	255	246	240	234	229	222	215	207	202	198	194	189	184	178	
23	428	342	303	280	264	253	244	237	232	227	220	213	205	200	196	191	186	181	175	
24	426	340	301	278	262	251	242	235	230	225	218	210	202	197	193	188	183	178	172	
25	424	339	299	276	260	249	240	234	228	224	216	209	201	196	192	187	182	177	171	
30	417	332	292	269	253	242	233	227	221	216	209	201	193	188	184	179	174	168	162	
40	408	323	284	261	245	234	225	218	212	208	200	192	184	178	174	169	164	158	151	
60	400	315	276	253	237	225	217	210	204	199	192	184	175	169	165	159	153	147	139	
120	392	307	268	245	229	217	209	202	196	191	183	175	166	160	155	150	143	135	125	
∞	384	300	260	237	221	210	201	194	188	183	175	167	157	151	146	139	132	122	100	

Table 6(b) Values of F_{α}

$\nu_1 =$ Degrees of Freedom for Denominator	$\nu_2 =$ Degrees of Freedom for Numerator																			
	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120	∞	
1	4052	5000	5403	5625	5764	5880	5928	5982	6033	6085	6137	6190	6240	6291	6341	6390	6439	6486	6536	
2	9850	9900	9917	9925	9930	9933	9936	9937	9939	9940	9942	9943	9945	9946	9947	9948	9949	9950	9950	
3	3412	3033	2646	2258	1870	1482	1094	706	318	277	275	273	272	271	270	269	268	267	266	
4	2120	1830	1606	1379	1152	925	698	471	244	203	201	200	199	198	197	196	195	194	193	
5	1620	1327	1208	1139	1097	1067	1046	1029	1016	1005	998	992	985	978	970	961	951	941	932	
6	1375	1092	978	915	875	847	826	810	798	788	772	766	760	750	743	734	724	714	706	
7	1225	925	845	785	746	719	699	684	672	662	647	641	631	621	611	601	591	582	574	
8	1126	865	759	701	663	637	618	603	591	581	567	561	551	541	531	521	511	501	491	
9	1056	802	699	642	606	580	561	547	535	525	511	505	495	485	475	465	455	445	431	
10	1004	756	655	599	564	539	520	506	494	485	471	465	455	445	435	425	415	408	391	
11	965	721	622	567	532	507	489	474	463	454	440	435	425	410	401	394	386	378	369	
12	933	693	595	541	506	482	464	449	430	420	406	401	391	382	376	370	362	354	345	
13	907	670	574	521	486	462	444	430	419	410	396	392	382	376	370	362	354	345	336	
14	886	651	556	504	469	446	428	414	403	394	380	376	366	361	351	341	335	327	318	
15	868	636	542	490	456	432	414	400	389	380	367	362	352	347	338	321	313	309	296	
16	853	623	529	477	444	420	403	389	378	369	365	341	336	316	310	302	293	284	275	
17	840	611	518	466	433	410	393	379	368	359	346	341	316	310	302	292	283	274	265	
18	829	601	508	456	423	401	384	371	360	351	337	332	308	298	284	275	266	257	248	
19	818	591	501	450	417	394	377	363	352	343	330	315	310	291	284	276	267	258	249	
20	810	585	494	443	410	387	370	356	346	337	323	309	294	284	278	269	261	252	242	
21	802	578	487	437	404	381	364	351	340	331	317	303	288	279	272	264	255	246	236	
22	795	572	482	431	399	376	359	345	335	326	312	298	283	273	267	258	250	240	231	
23	788	566	476	426	394	371	354	340	330	321	307	293	278	269	262	254	245	235	226	
24	782	561	472	422	390	367	350	336	326	317	303	289	274	264	258	249	240	231	221	
25	777	557	468	418	385	363	346	332	322	313	299	285	270	260	254	245	236	227	217	
30	756	539	451	402	370	347	330	317	307	298	284	270	255	245	239	230	221	211	201	
40	731	518	431	383	351	329	312	299	289	280	266	252	237	227	220	211	202	192	180	
60	708	498	413	365	334	312	295	282	272	263	250	235	220	210	203	194	184	174	160	
120	685	479	395	348	317	296	279	266	256	247	234	219	204	195	186	179	169	153	138	
∞	663	461	378	332	302	280	264	251	241	232	218	204	188	177	170	160	147	132	100	

Figure 9: F distribution

Solution. By Eq. (33)

$$F_{0.95}(10, 20) = \frac{1}{F_{0.05}(20, 10)}.$$

Using the F table (Figure 9) to get

$$\frac{1}{F_{0.05}(20, 10)} = \frac{1}{2.77} = 0.3610.$$

Therefore, $F_{0.95}(10, 20) = 0.3610$.

Example 17 Let t be distributed as a t -distribution with ν degrees of freedom.

(a) Use the representation of t to show that t^2 has an F distribution with $(1, \nu)$ degrees of freedom.

(b) Use part (a) to show that $t_{\alpha/2}^2 = F_{\alpha}(1, \nu)$.

Solution.

(a). By Eq. (30), assume t is in the form of

$$t = \frac{Z}{\sqrt{\frac{\chi^2(\nu)}{\nu}}},$$

where $Z \sim N(0, 1)$ and $\chi^2(\nu)$ has a chi square distribution of ν degrees of freedom. So

Z^2 has a chi square distribution of one degree of freedom. Consequently by Definition 2,

t^2 has an F distribution with $(1, \nu)$ degrees of freedom.

(b). Since

$$P(t > t_{\alpha/2}) = \frac{\alpha}{2},$$

we have (note that t -distribution is symmetric)

$$P(t^2 \leq t_{\alpha/2}^2) = 1 - \alpha.$$

On the other hand, by part (a), t^2 has an F distribution with $(1, \nu)$ degrees of freedom.

Therefore

$$P(t^2 \leq F_{\alpha}(1, \nu)) = 1 - P(t^2 \geq F_{\alpha}(1, \nu)) = 1 - \alpha.$$

Therefore, $t_{\alpha/2}^2 = F_{\alpha}(1, \nu)$.