9/25/04

Chi-square function

1. Using chi-square for goodness-of-fit testing from Zar's *Biostatistical Analysis*
2. Chi-square table from Rohlf & Sokal's *Statistical Tables*
3. Instructions on computing chi-square function in Python.

BIOSTATISTICAL ANALYSIS

FOURTH EDITION

JERROLD H. ZAR

Department of Biological Sciences
Northern Illinois University

PRENTICE HALL
Upper Saddle River, New Jersey 07458

22.1 CHI-SQUARE GOODNESS OF FIT

It is frequently desired to obtain a sample of nominal scale data and to infer whether the population from which it came conforms to a specified theoretical distribution. For example, a plant geneticist may raise 100 progeny from a cross that is hypothesized to result in a 3:1 phenotypic ratio of yellow-flowered to green-flowered plants. Perhaps a ratio of 84 yellow:16 green is observed, although out of this total of 100 plants, the geneticist's hypothesis would predict a ratio of 75 yellow:25 green. The question to be asked, then, is whether the observed frequencies (84 and 16) deviate significantly from the frequencies expected if the hypothesis were true (75 and 25).

The statistical procedure for attacking the question first involves the concise statement of the hypothesis to be tested. The hypothesis in this case is that the population which was sampled has a 3:1 ratio of yellow-flowered to green-flowered plants. As introduced in Section 6.4, this is referred to as a *null hypothesis* (abbreviated H_0), because it is a statement of "no difference"; in this instance, we are hypothesizing that the population flower color ratio is not different from 3:1. If it is concluded that H_0 is false, then an *alternate hypothesis* (abbreviated H_A) will be assumed to be true. In this case, H_A would be that the population sampled has a flower-color ratio which is *not* 3 yellow:1 green. Recall that one states a null hypothesis and an alternate hypothesis for every statistical test performed, and all possible outcomes are accounted for by the two hypotheses.

The following calculation of a statistic called *chi-square* is used as a measure of how far a sample distribution deviates from a theoretical distribution:

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i}. \qquad (22.1)^*$$

Here, f_i is the frequency, or number of counts, observed in class i, $\hat{f}_i$ is the frequency expected in class i if the null hypothesis is true,[†] and the summation is performed over all k categories of data; χ is lowercase Greek chi. Example 22.1 shows the chi-square calculation for the flower-color data presented. In this sample, there are two categories of data (i.e., $k = 2$): yellow-flowered plants and green-flowered plants. The expected frequency, $\hat{f}_i$, of each class is calculated by multiplying the total number of observations, n, by the proportion of the total that the null hypothesis predicts for the class. Therefore, for the two classes in the example, $\hat{f}_1 = (100)(\frac{3}{4}) = 75$ and $\hat{f}_2 = (100)(\frac{1}{4}) = 25$.

---

*Equation 22.1 can be rewritten as

$$\chi^2 = \sum_{i=1}^{k} \frac{f_i^2}{\hat{f}_i} - n, \qquad (22.2)$$

where n is the sum of all the f_i's, namely the total number of observations in the sample. Although this formula renders the calculation of χ^2 a little easier, it has the big disadvantage of not enabling us to examine each contribution to χ^2 [i.e., each $(f_i - \hat{f}_i)^2/\hat{f}_i$], and, as shown in Section 22.3, such an examination is an aid in determining how one might subdivide an overall chi-square analysis into component chi-square analyses for additional data collection. Thus, Equation 22.2 is seldom encountered.

†The symbol $\hat{f}$ is pronounced "f hat."

EXAMPLE 22.1 Calculation of chi-square goodness of fit, of data consisting of the colors of 100 flowers, to a hypothesized color ratio of 3:1.

H_0: The sample data came from a population having a 3:1 ratio of yellow to green flowers.

H_A: The sample data came from a population not having a 3:1 ratio of yellow to green flowers.

The data recorded are the 100 observed frequencies, f_i, in each of the two flower color categories, with the frequencies expected under the null hypothesis, $\hat{f}_i$, in parentheses.

	Category (flower color)		
	Yellow	*Green*	n
f_i	84	16	100
$(\hat{f}_i)$	*(75)*	*(25)*	

degrees of freedom $= \nu = k - 1 = 2 - 1 = 1$

$$\chi^2 = \sum \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i} = \frac{(84 - 75)^2}{75} + \frac{(16 - 25)^2}{25}$$

$$= \frac{9^2}{75} + \frac{9^2}{25}$$

$$= 1.080 + 3.240$$

$$= 4.320$$

$\chi^2_{0.05,1} = 3.841$

Therefore, reject H_0.

$0.025 < P < 0.05 \quad [P = 0.038]$

An improved procedure is presented in Section 22.4 (Example 22.4).

It should be apparent, by examining Equation 22.1, that larger disagreement between observed and expected frequencies (i.e., larger $f_i - \hat{f}_i$ values) will result in a larger χ^2 value. Thus, this type of calculation is referred to as a measure of *goodness of fit* (although it might better have been named a measure of "poorness of fit"). A calculated χ^2 value can be as small as zero, in the case of a perfect fit (i.e., each f_i value equals its corresponding $\hat{f}_i$), or very large if the fit is very bad; it can never be a negative value.

It is fundamentally important to appreciate that the chi-square statistic is calculated using the actual frequencies observed. It is not valid to convert the data to percentages and to attempt to submit the percentages to Equation 22.1. An additional consideration in calculating chi-square is described in Section 22.4.

What is meant by statistical significance in goodness of fit testing, as with other statistical tests, is derived from considerations of probability. Consider that the null hypothesis is true, i.e., the geneticist sampled a population of plants in which the yellow-to-green ratio is indeed 3 to 1. What we wish to ask is if it is likely to obtain from such a population a random sample of plants having an 84:16 flower-color ratio. If such a sample ratio can occur reasonably often, then we have no cause to reject H_0. If, however, there is little chance of obtaining a departure at least as great as 84 yellow-flowered and 16 green-flowered plants in a random sample from a population with a 3:1 ratio, then we may infer that the null hypothesis is false and that the alternate hypothesis is true (i.e., the sample came from a population with a color ratio that is not 3:1).

The computation of the probabilities that we require involves such complex mathematics that we are fortunate in having available tables of chi-square probabilities to aid in hypothesis testing. Appendix Table B.1 is a table of χ^2 values having certain probabilities of occurrence if H_0 is true.

By referring to the first line in the body of this table (the line for ν of 1), we can see, for example, that the probability ($P = \alpha$) of a χ^2 equal to or greater than 2.706 is 0.10 (i.e., 10%); this statement can be written concisely as $P(\chi^2 \geq 2.706) = 0.10$. As another example, we see that $P(\chi^2 \geq 3.841) = 0.05$. Now, in Example 22.1 we obtained $\chi^2 = 4.320$. By consulting Appendix Table B.1, we see that this value has associated with it a probability somewhere between 0.025 and 0.05, for $P(\chi^2 \geq 5.024) = 0.025$ and $P(\chi^2 \geq 3.841) = 0.05$. Thus, for this example, one can state that $0.025 < P(\chi^2 \geq 4.320) < 0.05$, or, simply, $0.025 < P < 0.05$. What this table tells us is that if H_0 were true, and if we repeated this same experiment a very large number of times, we could expect to get results that deviate at least this much from the hypothetical frequencies form 2.5% to 5% of the time.*

As explained in Section 6.4, biostatisticians often specify that if the magnitude of a calculated test statistic (such as χ^2) has an associated probability of 5% or less, its occurrence is so unlikely to be due to random sampling alone that we may reasonably conclude that the null hypothesis is false. This is the case for the data in Example 22.1; therefore, we reject H_0 and accept H_A concluding that the population sampled has a flower color ratio other than 3 yellow:1 green.

22.2 Chi-Square Goodness of Fit for more than Two Categories

Example 22.1 considered chi-square testing for goodness of fit when there are two categories of data (i.e., $k = 2$). This analysis may be extended readily to any larger number of classes, as Example 22.2 exemplifies. Here, 250 plants were examined ($n = 250$), seeds were classified into four categories ($k = 4$), and the calculated χ^2, using Equation 22.1, is 8.972. (We shall routinely express a calculated chi-square to three decimal places, because that is the accuracy of the table of critical values, Appendix Table B.1. Therefore, to avoid rounding error, we shall perform all intermediate computations, including those of $\hat{f}_i$, to four decimal places.)

*Some calculators and computer programs have the capability of determining the exact probability of a given χ^2 (e.g., Guenther, 1977). For the present example, we would thereby find that $P(\chi^2 \geq 4.320) = 0.038$.

EXAMPLE 22.2 Chi-square goodness of fit for $k = 4$.

H_0: The sample comes from a population having a 9:3:3:1 ratio of yellow-smooth to yellow-wrinkled to green-smooth to green-wrinkled seeds.

H_A: The sample comes from a population not having a 9:3:3:1 ratio of the above four seed phenotypes.

The sample data are recorded as observed frequencies, f_i, with the frequencies expected under the null hypothesis, $\hat{f}_i$ in parenthesis.

	Yellow smooth	*Yellow wrinkled*	*Green smooth*	*Green wrinkled*	*n*
f_i	152	39	53	6	250
$(\hat{f}_i)$	*(140.6250)*	*(46.8750)*	*(46.8750)*	*(15.26250)*	

$\nu = k - 1 = 3$

$$\chi^2 = \frac{11.3750^2}{140.6250} + \frac{7.8750^2}{46.8750} + \frac{6.1250^2}{46.8750} + \frac{9.6250^2}{15.6250}$$
$$= 0.9201 + 1.3230 + 0.8003 + 5.9290$$
$$= 8.972$$

$\chi^2_{0.05,3} = 7.815$

Therefore, reject H_0.

$0.025 < P < 0.05 \quad [P = 0.030]$

It has already been pointed out that larger χ^2 values will result from larger differences between f_i and $\hat{f}_i$, but large calculated χ^2 values may also simply be the result of a large number of classes of data, because the calculation involves the summing over all classes. Thus, in considering the significance of a calculated χ^2, the value of k must in some way be taken into account. What is done is to consider the quantity known as *degrees of freedom** (abbreviated DF, or by lowercase Greek nu, ν). For the goodness of fit testing discussed in this chapter, DF (i.e., ν) $= k - 1$. Thus, for Example 22.2, DF $= 4 - 1 = 3$, while the calculated $\chi^2 = 8.972$. Entering Appendix Table B.1 in the row for 3 DF, we see that $P(\chi^2 \geq 7.815) = 0.05$, and $P(\chi^2 \geq 9.348) = 0.025$. Therefore, $0.025 < P(\chi^2 \geq 8.972) < 0.05$, and we would reject the null hypothesis which states that the sample came from a population having a 9:3:3:1 phenotypic ratio of yellow-smooth:yellow-wrinkled:green-smooth:green-wrinkled seeds. Tabled critical values are frequently denoted as $\chi^2_{\alpha,\nu}$, so we could write $\chi^2_{0.05,3} = 7.815$, $\chi^2_{0.10,3} = 6.251$, etc.

When we say, in a goodness of fit problem such as Example 22.1 or 22.2, that DF $= k - 1$, we are stating that, given the frequencies in any $k - 1$ of the categories, we

*This term was introduced by R. A. Fisher, in 1922, while discussing contingency tables (see Chapter 23) (David, 1995).

can readily calculate the frequency in the remaining category. This is true because n is known, and the sum of the frequencies in all k categories equals n. (In other words, one has "freedom" in assigning frequencies to only $k-1$ of the categories.) Another way of looking at chi-square degrees of freedom is to note that DF equals k minus the number of sample constants used to calculate the expected frequencies. In the present examples, only one constant, n, was so used, so $\nu = k - 1$.

22.3 Subdividing Chi-Square Analyses

In Example 22.2, the chi-square analysis detected a difference between the observed and expected frequencies too great to be attributed to chance, and the null hypothesis was rejected. This conclusion may be satisfactory in some instances, but in many cases the investigator will wish to perform further analysis.

For the example under consideration, the null hypothesis is that the sample came from a population having a 9:3:3:1 phenotypic ratio. If the chi-square analysis had not led to a rejection of the hypothesis, we would proceed no further. But since H_0 was rejected, we may wish to ask whether the significant disagreement between observed and expected frequencies was concentrated in certain of the classes, or whether the difference was due to the effects of the data in all of the classes. Of the four individual contributions to the chi-square value—0.9201, 1.3230, 0.8003, and 5.9290—that resulting from the last class (the green-wrinkled seeds) contributes a relatively large amount to the size of the calculated χ^2. Thus we see that the nonconformity of the sample frequencies to those expected from a population with a 9:3:3:1 ratio is due largely to the magnitude of the discrepancy between f_4 and $\hat{f}_4$.

This line of thought can be examined as shown in Example 22.3. First, we test H_0: f_1, f_2, and f_3 came from a population having a 9:3:3 ratio. (H_A: The frequencies in the first three categories came from a population having a phenotypic ratio other than 9:3:3.) This null hypothesis is not rejected, indicating that the frequencies in the first

EXAMPLE 22.3 Chi-square goodness of fit, subdividing the chi-square analysis of Example 22.2.

H_0: The sample came from a population with a 9:3:3 ratio of the first three phenotypes in Example 22.2.

H_A: The sample came from a population not having a 9:3:3 ratio of the first three phenotypes in Example 22.2.

Seed Characteristics

	Yellow smooth	*Yellow wrinkled*	*Green smooth*	*n*
f_i	152	39	53	244
$(\hat{f}_i)$	*(146.4000)*	*(48.80000)*	*(48.80000)*	

22.5 Bias in Chi-Square Calculations

In order for us to assign a probability to the results of a chi-square goodness of fit test, and thereby assess the statistical significance of the test, the calculated χ^2 must be a close approximation to the theoretical distribution that is summarized in Appendix Table B.1. This approximation is quite acceptable as long as the expected frequencies are not too small. If $\hat{f}_i$ values are very small, however, the calculated χ^2 is biased in that it is larger than the theoretical χ^2 it is supposed to estimate, and there is a tendency to reject the null hypothesis with a probability greater than α. This is clearly undesirable, and for decades statisticians have attempted to define in a convenient manner what would constitute $\hat{f}_i$'s that are "too small."

For decades a commonly applied general rule was that no expected frequency should be less than 5.0*, even though it has long been known that is it tolerable to have a few $\hat{f}_i$'s considerably smaller than that (e.g., Cochran, 1952, 1954). By a review of previous recommendations and an extensive empirical analysis, Roscoe and Byars (1971) reached conclusions that provide less restrictive guidelines for chi-square goodness of fit testing. They found that the test is remarkably robust when testing for a uniform distribution—i.e., for H_0: In the population, the frequencies in all k categories are equal—in which case $\hat{f}_i = n/k$. In this situation, it appears that it is acceptable to have expected frequencies as small as 1.0 for testing at α as small as 0.05, or as small as 2.0 for α as small as 0.01. The chi-square test works nearly as well when there is moderate departure from a uniform distribution in H_0, and the average expected frequencies may be as small as those indicated for a uniform distribution. And even with great departure from uniform it appears that the average expected frequency (i.e., n/k) may be as low as 2.0 for testing at α as low as 0.05 and as low as 4.0 for α as small as 0.01. Koehler and Larntz (1980) suggest that the chi-square test is applicable for situations where $k \geq 3, n \geq 10$, and $n^2/k \geq 10$. Users of goodness of fit testing can be comfortable if their data fit both the Roscoe and Byars and the Koehler and Larntz guidelines. These recommendations are for situations where there are more than 2 categories. If $k = 2$, then it is wise to have $\hat{f}_i$'s of at least 5.0, or to use the binomial test as indicated in the next paragraph.

The chi-square calculation can be made if the data for the classes with the offensively low $\hat{f}_i$ values are simply eliminated from H_0 and the subsequent analysis. Or, certain of the classes of data might be meaningfully combined so as to result in all $\hat{f}_i$ values being large enough to proceed with the analysis. Such modified procedures are not to be recommended as routine practice. Rather, the experimenter should strive to obtain a sufficiently large n for the analysis to be performed. When $k = 2$ and each f_i is small, the use of the binomial test (Section 24.6) is preferable to chi-square analysis. [Similarly, use of the multinomial, rather than the binomial, distribution is appropriate when $k > 2$ and the f_i's are small; however, this is a tedious procedure and will not be demonstrated here (Radlow and Alf, 1975).]

*Some statisticians were even stricter, recommending a lower limit of 10.0.

(2)

Statistical Tables
THIRD EDITION

F. James Rohlf
Robert R. Sokal

State University of New York
at Stony Brook

W. H. Freeman and Company
New York

TABLE **D** Critical values of the chi-square distribution

This table furnishes critical values for the chi-square distribution for degrees of freedom $\nu = 1$ to 100 in increments of one. The percentage points (corresponding to $\alpha = 0.995$, 0.975, 0.9, 0.5, 0.1, 0.05, 0.025, 0.01, 0.005, and 0.001) represent the area to the right of the critical value of χ^2 in one tail of the distribution, as shown in the accompanying figure. The critical values of χ^2 are given to three significant decimal places, except when $\chi^2 > 100$, in which case they are given to two significant decimal places.

To find the critical value of χ^2 for a given number of degrees of freedom, look up ν in the left (argument) column of the table and read off the desired values of χ^2 in that row. For example, for 8 degrees of freedom, $\chi^2_{.05[8]} = 15.507$ and $\chi^2_{.01[8]} = 20.090$. The last value indicates that 1% of the area of the chi-square distribution (for 8 degrees of freedom) is to the right of the value of $\chi^2 = 20.090$. For values of $\nu > 100$, compute approximate critical values of χ^2 using the following formula: $\chi^2_{\alpha[\nu]} = \frac{1}{2}(t_{2\alpha[\infty]} + \sqrt{2\nu - 1})^2$, where $t_{2\alpha[\infty]}$ comes from Table **B.** Thus $\chi^2_{.05[120]}$ is computed as $\frac{1}{2}(t_{.10[\infty]} + \sqrt{240 - 1})^2 = \frac{1}{2}(1.645 + \sqrt{239})^2 = \frac{1}{2}(17.10462)^2 = 146.284$. For $\alpha > 0.5$, employ $-t_{2(1-\alpha)[\infty]}$ in this formula. When $\alpha = 0.5$, $t_{2\alpha} = 0$.

There are numerous applications of the chi-square distribution (Section 7.6) in statistics. It is used to set confidence limits to sample variances (Section 7.7), to test the homogeneity of variances (Section 13.3) and of correlation coefficients (Section 15.5), and to test goodness of fit (Chapter 17).

Except for $\alpha = 0.001$, the values of chi-square from 1 to 30 degrees of freedom have been taken from a more extensive table by C. M. Thompson (*Biometrika* **32:**188–189, 1941) with permission of the publisher. Values between 31 and 100 degrees of freedom were approximated using the Cornish–Fisher asymptotic expansion following the account of M. Zelen and N. C. Severo, section 26.2.49 of M. Abramovitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions* (U.S. National Bureau of Standards, 1964). The values of the Hermite polynomials were computed, instead of using the tables in the source above. All values for $\alpha = 0.001$ were taken from H. L. Harter, *New Tables of the Incomplete Gamma-Function Ratio and of Percentage Points of the Chi-Square and Beta Distribution* (U.S. Government Printing Office, 1964).

TABLE D Critical values of the chi-square distribution

ν \ α	.995	.975	.9	.5	.1	.05	.025	.01	.005	.001	α / ν
1	0.000	0.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	10.828	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	13.816	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	16.266	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	18.467	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	20.515	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	22.458	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	24.322	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	26.124	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	27.877	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	29.588	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	31.264	11
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	32.910	12
13	3.565	5.009	7.042	12.340	19.812	22.362	24.736	27.688	29.819	34.528	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	36.123	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	37.697	15
16	5.142	6.908	9.312	15.338	23.542	26.296	28.845	32.000	34.267	39.252	16
17	5.697	7.564	10.085	16.338	24.769	27.587	30.191	33.409	35.718	40.790	17
18	6.265	8.231	10.865	17.338	25.989	28.869	31.526	34.805	37.156	42.312	18
19	6.844	8.907	11.651	18.338	27.204	30.144	32.852	36.191	38.582	43.820	19
20	7.434	9.591	12.443	19.337	28.412	31.410	34.170	37.566	39.997	45.315	20
21	8.034	10.283	13.240	20.337	29.615	32.670	35.479	38.932	41.401	46.797	21
22	8.643	10.982	14.042	21.337	30.813	33.924	36.781	40.289	42.796	48.268	22
23	9.260	11.688	14.848	22.337	32.007	35.172	38.076	41.638	44.181	49.728	23
24	9.886	12.401	15.659	23.337	33.196	36.415	39.364	42.980	45.558	51.179	24
25	10.520	13.120	16.473	24.337	34.382	37.652	40.646	44.314	46.928	52.620	25
26	11.160	13.844	17.292	25.336	35.563	38.885	41.923	45.642	48.290	54.052	26
27	11.808	14.573	18.114	26.336	36.741	40.113	43.194	46.963	49.645	55.476	27
28	12.461	15.308	18.939	27.336	37.916	41.337	44.461	48.278	50.993	56.892	28
29	13.121	16.047	19.768	28.336	39.088	42.557	45.722	49.588	52.336	58.301	29
30	13.787	16.791	20.599	29.336	40.256	43.773	46.979	50.892	53.672	59.703	30
31	14.458	17.539	21.434	30.336	41.422	44.985	48.232	52.191	55.003	61.098	31
32	15.134	18.291	22.271	31.336	42.585	46.194	49.480	53.486	56.329	62.487	32
33	15.815	19.047	23.110	32.336	43.745	47.400	50.725	54.776	57.649	63.870	33
34	16.501	19.806	23.952	33.336	44.903	48.602	51.966	56.061	58.964	65.247	34
35	17.192	20.569	24.797	34.336	46.059	49.802	53.203	57.342	60.275	66.619	35
36	17.887	21.336	25.643	35.336	47.212	50.998	54.437	58.619	61.582	67.985	36
37	18.586	22.106	26.492	36.335	48.363	52.192	55.668	59.892	62.884	69.346	37
38	19.289	22.878	27.343	37.335	49.513	53.384	56.896	61.162	64.182	70.703	38
39	19.996	23.654	28.196	38.335	50.660	54.572	58.120	62.428	65.476	72.055	39
40	20.707	24.433	29.051	39.335	51.805	55.758	59.342	63.691	66.766	73.402	40
41	21.421	25.215	29.907	40.335	52.949	56.942	60.561	64.950	68.053	74.745	41
42	22.138	25.999	30.765	41.335	54.090	58.124	61.777	66.206	69.336	76.084	42
43	22.859	26.785	31.625	42.335	55.230	59.304	62.990	67.459	70.616	77.419	43
44	23.584	27.575	32.487	43.335	56.369	60.481	64.202	68.710	71.893	78.750	44
45	24.311	28.366	33.350	44.335	57.505	61.656	65.410	69.957	73.166	80.077	45
46	25.042	29.160	34.215	45.335	58.641	62.830	66.617	71.201	74.437	81.400	46
47	25.775	29.956	35.081	46.335	59.774	64.001	67.821	72.443	75.704	82.720	47
48	26.511	30.755	35.949	47.335	60.907	65.171	69.023	73.683	76.969	84.037	48
49	27.249	31.555	36.818	48.335	62.038	66.339	70.222	74.919	78.231	85.351	49
50	27.991	32.357	37.689	49.335	63.167	67.505	71.420	76.154	79.490	86.661	50

TABLE D Cri

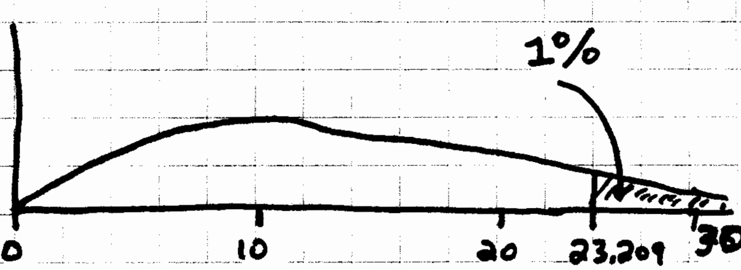
ν \ α	.995	.975
51	28.735	33.162
52	29.481	33.96[illegible]
53	30.230	34.776
54	30.981	35.58[illegible]
55	31.735	36.39[illegible]
56	32.490	37.21[illegible]
57	33.248	38.027
58	34.008	38.84[illegible]
59	34.770	39.662
60	35.534	40.482
61	36.300	41.303
62	37.068	42.126
63	37.838	42.950
64	38.610	43.776
65	39.383	44.603
66	40.158	45.431
67	40.935	46.261
68	41.713	47.092
69	42.494	47.924
70	43.275	48.758
71	44.058	49.592
72	44.843	50.428
73	45.629	51.265
74	46.417	52.103
75	47.206	52.942
76	47.997	53.782
77	48.788	54.623
78	49.582	55.466
79	50.376	56.309
80	51.172	57.153
81	51.969	57.998
82	52.767	58.845
83	53.567	59.692
84	54.368	60.540
85	55.170	61.389
86	55.973	62.239
87	56.777	63.089
88	57.582	63.941
89	58.389	64.793
90	59.196	65.647
91	60.005	66.501
92	60.815	67.356
93	61.625	68.211
94	62.437	69.068
95	63.250	69.925
96	64.063	70.783
97	64.878	71.642
98	65.694	72.501
99	66.510	73.361
100	67.328	74.222

9/25/04

Computing chi-square function in Python:

- Download package SciPy from www.scipy.org/download and install.
- In Python:

```
>>> import scipy.special
>>> nu = 10            #degrees of freedom
>>> p = 0.01           # cutoff probability (alpha)
>>> scipy.special.chdtri(nu,p)
23.209251158954359
```



- The scipy package has many functions & sub-packages; documentation is available at www.scipy.org or by entering (in Python)

```
help(scipy)            #after importing scipy
help(scipy.special)
etc...
```