5.1 The χ^2 test

Suppose we want to test a theory that predicts a certain value of a quantity A. Suppose the predicted (i.e., expected) value is E. Here the null hypothesis is that E is truly the value of the parameter A. The alternative hypothesis is that it is not.

You have performed an experiment to measure that value, and you have got an observed value O. The obtained value O would never be exactly the same as E. Still we would like to know how good this fit is. Does the experimental result support the theory, even though the experimental result did not exactly match the expectation from theory? If the difference between O and E behaves just as a random variable, we can assume the difference to have occurred due to random chance. But if it differs significantly from the behavior of a random variable, we have to admit that there may be a causal factor behind it, namely, that the experiment is not conforming to the theory.

A good way of judging this would be to obtain the square of the difference $(O-E)^2$, which will be a positive number, and to normalize it by dividing it by the expected value E. Let us call this result X^2. Thus,

$$X^2 = \frac{(O-E)^2}{E}.$$

The question is, what is the probability of getting such a value, if such "squared differences" are distributed as a random variable?

To decide this issue, we have to compare it with the character of a pure random variable. Consider a random variable Y that is distributed as per a normal distribution with mean 0 and variance 1. Now define a new variable Q which is the square of Y, i.e., $Q = Y^2$. What will be the probability deinsity function of the variable Q? Firstly, since we are taking a square, the values cannot be negative. Secondly, since it follows a normal distribution, a large number of samples of Y will be will be crowded around zero. Squaring these will yield even smaller numbers. Thus, it will have a very large frequency close to zero, which will fall off

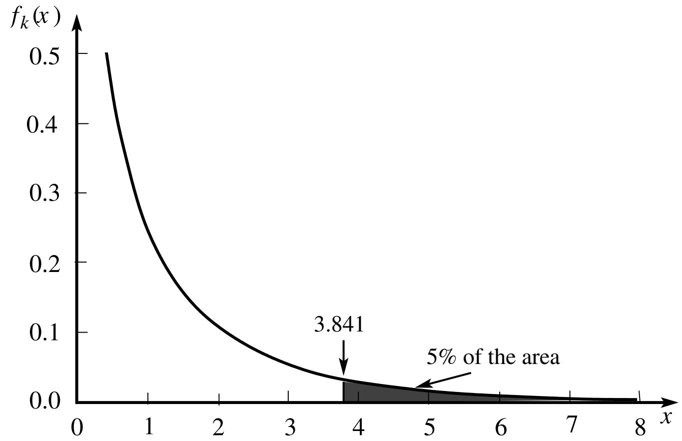


Figure 5.1: The χ^2 distribution for one normal distributed variable.

for larger values of Q, as shown in Fig. 5.1. Such a distribution is called a χ^2 distribution (pronounced as 'kai-square'). Thus Q is a χ^2 distributed random variable.

This graph allows us to figure out what is the probability of finding a χ^2 value greater than a given value. This is the area under the curve to the right of that value. For example, for the curve in Fig. 5.1, 5% of the area lies above $\chi^2 = 3.841$. This means that the probability of getting a value greater than 3.841 is only 0.05. We now compare it with the value of X^2 obtained in the experiment. If such a large value of χ^2 is obtained in the experiment, we can assume that such a deviation between the expected value and the observed value is unlikely to occur due to random error.

We have so far come across the concept of 'level of confidence' (C). Its dual is the 'level of significance' (S), which is nothing but $S = 1 - C$. Thus, a 0.05 level of significance is the same as 0.95 (or 95%) level of confidence. Since 5% of the area lies beyond $\chi^2 = 3.841$, we can reject the null hypothesis with 95% confidence if the obtained value of X^2 lies beyond this value.

A major advantage of the χ^2 test is that we need not just

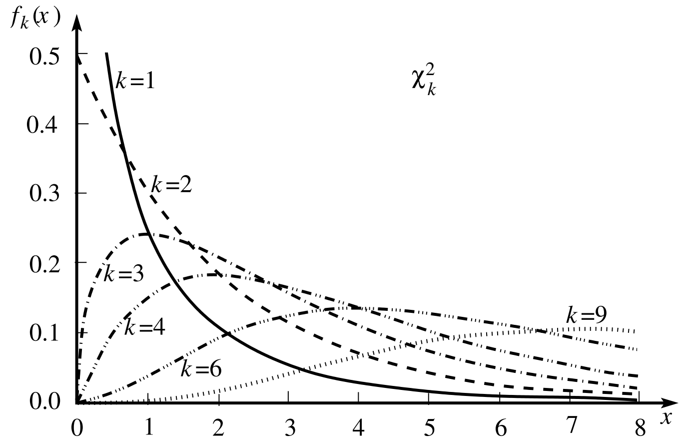


Figure 5.2: The graphs of χ^2 distributions for different degrees of freedom.

compare one expected value with one obtained value. We can compare a range of such values.

For that, we first investigate the probability density function of some other random variables. We now take two independent random variables Y_1 and Y_2, both having normal distribution with mean zero and standard deviation 1. Then we define a new random variable as $Q_2 = Y_1^2 + Y_2^2$. The subscript of Q shows what is known as the 'degree of freedom' or the number of random variables taken to construct the variable Q. In that sense the Q constructed out of just one random variable Y should be termed as Q_1. That way we can construct a series of random variables as

$$
\begin{aligned}
Q_1 &= Y_1^2 & k &= 1 \\
Q_2 &= Y_1^2 + Y_2^2 & k &= 2 \\
Q_3 &= Y_1^2 + Y_2^2 + Y_3^2 & k &= 3 \\
&\vdots & &\vdots \\
Q_n &= Y_1^2 + Y_2^2 + Y_3^2 + \cdots + Y_n^2 & k &= n
\end{aligned}
$$

Here k denotes the degree of freedom. The probability density functions of some of these random variables are shown in Fig 5.2.

Now, depending on the experimental situation, the test statistic is calculated as

$$X^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

This is then compared with the χ^2 values using the table given in Table 5.4.

The are a few conditions of applicability of the χ^2 test. First, the sampling should be truly random, i.e., without any bias in favour or against certain characteristics. Second, the *expected* number in each category should be at least 5. Third, the sample size should be no more than 10% the population size.

Let us illustrate this with examples.

Example 5.1: Some characteristics in pea plants are typically used in genetics experiments, e.g., plants can be tall or short, the peas can be wrinkled or round, etc. Suppose we do not know if these characters are related or unrelated. We know from theory of genetics that if these traits are independently assorted, then if tall plants with round seeds are crossed with short plants with wrinkled seeds, then in the first generation all plants will be tall, with round seeds. But in the second generation these characteristics will show up in a 9:3:3:1 ratio, as shown in table 5.1.

Table 5.1: The values related to the example.

Type	Expected ratio	observed number
Tall, Round	9	84
Tall, wrinkled	3	32
short, Round	3	38
short, wrinkled	1	6

An experiment was done and 160 plants were studied. The results obtained are also listed in the third column of table 5.1.

Do the results support the hypothesis that these traits are independently assorted?

Solution:

Null hypothesis: The traits are independently assorted (the traits are distributed as 9:3:3:1)

Alternative hypothesis: The traits are not independently assorted (the above ratios do not hold)

First we work out how many of each category will be expected if the null hypothesis were true. We then calculate the value of $(O-E)^2/E$ for each row and add them up. The results are listed in Table 5.2.

Table 5.2: Computation of the test statistic for the example.

Type	Expected (E)	Observed (O)	$(O-E)^2/E$
Tall, Round	90	84	$(-6)^2/90$
Tall, wrinkled	30	34	$4^2/30$
short, Round	30	38	$8^2/30$
short, wrinkled	10	4	$(-6)^2/10$

This gives

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} = 6.66$$

Now we need to find the critical value of χ^2 beyond which the null hypothesis can be rejected with 95% confidence, i.e., (1-0.95)=0.05 level of significance. What is the degree of freedom in this case? Note that if the total number of data points (in this case, 160) is known, knowing any three of the observed numbers would enable anyone of determine the fourth. There is no freedom in choosing the 4th data point. Thus, the degree of freedom is $(4-1)=3$.

From Table 5.4 we find that for 3 degrees of freedom (3rd row), the critical value of χ^2 for 0.05 level of significance is 7.815. Since the obtained value of χ^2 is below this value, the null hypothesis cannot be rejected. □

Example 5.2: There is a commonly held belief that one's IQ (intelligence quotient) measured at a young age is a good indicator of his or her intellectual ability and hence promise of success in life. Suppose you set out to test the belief. You gather a set of a hundred 15-year-old kids coming from families of similar social status and education. You get their IQ tested, and divide them into three groups: those with IQ<80, those with 80≤IQ≤100 and those with IQ>100. Then you wait for 15 years and check if they have distinguished themselves in any avenue of intellectual activity. You again divide them into two groups depending on whether the answer is 'yes' or 'no'. The results were as follows. Out of the 16 kids who scored IQ<100, 12 showed intellectual success, and out of the 18 kids who scored IQ>100, 11 distinguished themselves in intellectual sphere. State the null and alternative hypotheses, and your conclusion out of the test.

Solution: The two hypotheses are:

H_0: Probability of intellectual success is independent of the measured IQ

H_1: Probability of intellectual success is dependent on the measured IQ

Note that the null hypothesis should always be of 'equality' type, and the alternative hypothesis should point to an inequality.

The contingency table obtained from the data is given in Table 5.3.

Out of a total of 100 people, 32 saw success. So in the whole population, the probability of intellectual success is 0.32. If the null hypothesis is true, the probability of intellectual success should be the same for all three categories. Therefore, among

Table 5.3: The contingency table pertaining to the example.

Category	Intellectual success		
	Yes	No	Total
IQ<80	4	12	16
Expected	5.12	10.88	
80≤IQ≤100	19	47	66
Expected	21.12	44.88	
IQ>100	9	9	18
Expected	5.76	12.24	
Total	32	68	100

the 14 people who scored below 80, we would expect $16 \times 0.32 = 5.12$ to succeed. Similarly we calculate the expected number of intellectual success among the other categories: $66 \times 0.32 = 21.12$ and $18 \times 0.32 = 5.76$. These expected values are then put into the contingency table.

The expected values of intellectual failure will be just the total number minus the number that are expected to be intellectually successful. The elements in each row must add up to the last entry in that row, and the elements of each column must add up to the last entry of the column. That way all the elements are calculated.

Now the χ^2 value is calculated as

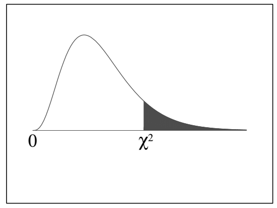
$$\begin{aligned}
\chi^2 &= \frac{(4-5.12)^2}{5.12} + \frac{(12-10.88)^2}{10.88} + \frac{(19-21.12)^2}{21.12} \\
&\quad + \frac{(47-44.88)^2}{44.88} + \frac{(9-5.76)^2}{5.76} + \frac{(9-12.24)^2}{12.24} \\
&= 3.353
\end{aligned}$$

Now we will have to refer to the χ^2 table. What will be the degree of freedom? Note that in each column there are 3 entries, but if we know two of them we can determine the third. There are two columns but it suffices to have information about one of them. Thus the degree of freedom is $2 \times 1 = 2$. In general, if there are n columns and m rows, then the degree of freedom is $(n-1)(m-1)$.

We find from the table that for 2 degrees of freedom, the value of χ^2 has to be at least 4.605 to reject the null hypothesis with significance level 0.1, i.e., 90% confidence. Since the value of χ^2 is below that, we cannot reject the null hypothesis and have to conclude that the IQ level is not a significant determinant of intellectual promise.

Notice that just a look at the observed data may induce one to believe that those with higher IQ are more likely to find intellectual success in later life, but the actual test points to the fact that this was only a common-sense conclusion obtained from a sample. If more samples are drawn and the test repeated, the bias will disappear on the average. $\square$

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Table 5.4: The table of χ^2 values for different degrees of freedom.