LINEAR RESTRICTIONS ON CHI-SQUARE

By Franklin E. Satterthwaite

University of Iowa

Chi-square is a statistic widely used in statistical analysis. It is usually of the form,

$$\begin{aligned} \chi^2 &= \textstyle\sum_1^n \chi_i^2 \\ &= \textstyle\sum_1^n \left(\frac{x_i - m_i}{\sigma_i}\right)^2, \end{aligned} \tag{1}$$

where the x_i's are independent normally distributed variables drawn from populations with respective means and standard deviations, m_i and σ_i. In practical problems the independence of the x_i's is often modified by placing restrictions on the χ_i's in order to estimate the m_i's or σ_i's. It is well known that if m such restrictions which are linear and homogeneous (also algebraically independent) are placed on the χ_i's, then the resulting chi-square, (1), is distributed according to the chi-square distribution with $n - m$ degrees of freedom. The purpose of this paper is to study the case where the restrictions are not necessarily homogeneous.

1. Geometrical development. The χ_i's of equation (1) may be considered as co-ordinates in an n-dimensional space. Equation (1) represents a sphere in such a space with its center at the origin and with radius, χ. We should like to determine the distribution of χ^2. First, since the χ_i's are independent, we may form their joint distribution,[1]

$$\begin{aligned} F(\chi_1, \chi_2, \cdots \chi_n)\, dV &= K\Pi_j e^{-\frac{1}{2}\chi_j^2}\, d\chi_j \\ &= Ke^{-\frac{1}{2}\Sigma\chi_j^2}\, d\chi_1\, d\chi_2 \cdots d\chi_n \\ &= Ke^{-\frac{1}{2}\chi^2}\, dV. \end{aligned} \tag{2}$$

We may change the variable in (2) to χ^2 if we can determine dV. Since the n-dimensional sphere represented by equation (1) has a volume proportional to χ^n, we may write

$$\begin{aligned} dV &= Kd(\chi^2)^{\frac{1}{2}n} \\ &= K(\chi^2)^{\frac{1}{2}n-1}\, d\chi^2. \end{aligned}$$

Substituting this value in the distribution (2) we obtain for the distribution of chi-square,

$$F(\chi^2)\, d\chi^2 = K(\chi^2)^{\frac{1}{2}n-1}e^{-\frac{1}{2}\chi^2}\, d\chi^2,$$

which is the usual form of the chi-square distribution for n degrees of freedom.

[1] The letter K will be used throughout as a constant, not necessarily the same constant from equation to equation.

We shall next restrict the values of χ_i by means of a condition,

$$a_{11}\chi_1 + a_{12}\chi_2 + \cdots + a_{1n}\chi_n = \rho_1, \qquad \Sigma a_{1j}^2 = 1, \tag{3}$$

where ρ_1 is a constant. This restriction represents a hyper-plane in our n-dimensional space at a distance ρ_1 from the origin. The intersection of this hyper-plane with our sphere (1) is an $(n - 1)$-dimensional sphere with radius

$$\chi' = (\chi^2 - \rho_1^2)^{\frac{1}{2}}.$$

The differential of the volume of this sphere is

$$dV = K(\chi^2 - \rho_1^2)^{\frac{1}{2}(n-1)-1}\, d\chi^2.$$

Substituting this in the distribution (2) we obtain the distribution of chi-square subject to the single linear restriction, (3). Thus

$$F(\chi^2)\, d\chi^2 = K(\chi^2 - \rho_1^2)^{\frac{1}{2}(n-1)-1} e^{-\frac{1}{2}\chi^2}\, d\chi^2,$$

or more conveniently,

$$F(\chi^2 - \rho_1^2)\, d(\chi^2 - \rho_1^2) = K(\chi^2 - \rho_1^2)^{\frac{1}{2}(n-1)-1} e^{-\frac{1}{2}(\chi^2-\rho_1^2)}\, d(\chi^2 - \rho_1^2).$$

The argument may be readily extended to include additional linear restrictions of the form,

$$\begin{aligned} a_{21}\chi_1 + a_{22}\chi_2 + \cdots + a_{2n}\chi_n &= \rho_2, \qquad \Sigma a_{2j}^2 = 1, \\ &\cdots \\ a_{m1}\chi_1 + a_{m2}\chi_2 + \cdots + a_{mn}\chi_n &= \rho_m, \qquad \Sigma a_{mj}^2 = 1. \end{aligned} \tag{4}$$

For convenience we shall assume that the restrictions form an orthogonal set[2] so that

$$\Sigma_j a_{ij} a_{kj} = 0, \qquad i \neq k.$$

The hyper-plane represented by equation (4) is at a distance, ρ_2, from the origin. Since (4) is orthogonal to (3), it is also at a distance, ρ_2, from the center of the $(n - 1)$-dimensional sphere obtained on applying the first restriction. Therefore the intersection of this hyper-plane with the $(n - 1)$-dimensional sphere will give an $(n - 2)$-dimensional sphere of radius

$$\chi'' = (\chi^2 - \rho_1^2 - \rho_2^2)^{\frac{1}{2}}.$$

Similarly, if we consider all m restrictions, we obtain an $(n - m)$-dimensional sphere with radius

$$\chi^{(m)} = (\chi^2 - \Sigma \rho_j^2)^{\frac{1}{2}}.$$

[2] Any set of linear restrictions which are algebraically independent and consistent may be replaced by an orthogonal set. Thus if (4) were not orthogonal to (3), we could replace (4) by (4) $- k$(3) where k is determined by the condition

$$\Sigma a_{1j}(a_{2j} - k a_{1j}) = 0$$

or

$$\Sigma a_{1j} a_{2j} = k \Sigma a_{1j}^2$$

The differential of the volume of this sphere will be

$$dV = K(\chi^2 - \Sigma\rho_j^2)^{\frac{1}{2}(n-m)-1} d(\chi^2 - \Sigma\rho_j^2).$$

Substituting this in (2) we see that

$$(\chi^{(m)})^2 = \chi^2 - \sum \rho_j^2$$

is distributed as is chi-square with $n - m$ degrees of freedom.

2. Alternate analytic development. It is perhaps desirable that we present an analytic proof of the foregoing theorem. Therefore we shall first regard the ρ_j's as variables and shall determine the joint distribution of χ^2 and the ρ_j's. We may then pass to the distribution of those values of χ^2 which correspond to assigned values of the ρ_j's. Note that the χ_i's are considered to be statistically independent.

The characteristic function of the joint distribution of χ^2 and the ρ_j's is known to be[3]

$$\phi(t, t_1, \cdots t_m) = \frac{e^{-Q/2(1-2it)}}{(1 - 2it)^{\frac{1}{2}n}},$$

where

$$Q = \sum_{i,j,k} a_{ik} a_{jk} t_i t_j$$
$$= \sum t_i^2, \qquad \text{since } \sum a_{ik} a_{jk} = \delta_{ij}.$$

Applying the Fourier transform, we obtain the joint distribution of χ^2 and the ρ_j's:

$$F(\chi^2, \rho_1, \cdots \rho_n) = K \int \cdots \int \frac{e^{Q'}}{(1 - 2it)^{\frac{1}{2}n}} dt_m \cdots dt_1 \, dt,$$

where

$$Q' = - it\chi^2 - \Sigma it_j \rho_j - \{\Sigma t_j^2 / 2(1 - 2it)\}$$
$$= - it\chi^2 - \frac{\Sigma[t_j + i\rho_j(1 - 2it)]^2}{2(1 - 2it)} - \tfrac{1}{2}(1 - 2it)\Sigma\rho_j^2.$$

Performing the integration with respect to $t_1, \cdots t_m$, we have,

$$F = Ke^{-\frac{1}{2}\Sigma\rho_j^2} \int \frac{e^{-it\chi^2}}{(1 - 2it)^{\frac{1}{2}n}} (1 - 2it)^{\frac{1}{2}m} e^{it\Sigma\rho_j^2} dt,$$

and finally,

$$F = K(\chi^2 - \Sigma\rho_j^2)^{\frac{1}{2}(n-m)-1} e^{-\frac{1}{2}\chi^2}.$$

[3] See A. T. Craig, "A certain mean value problem in statistics," *Bull. Amer. Math. Soc.*, Vol. 42 (1936), p. 671.

In our problem we want the distribution of χ^2 (or more conveniently, of $\chi^2 - \Sigma\rho_j^2$) when the ρ_j's take on fixed values. To obtain this we substitute fixed values, $\hat{\rho}_j$'s, into the joint distribution and divide by the marginal total,

$$\int F(\chi^2, \hat{\rho}_1, \hat{\rho}_2 \cdots \hat{\rho}_m)\, d\chi^2 = K\Gamma\left[\tfrac{1}{2}(n-m)\right]2^{\frac{1}{2}(n-m)}\, e^{-\frac{1}{2}\Sigma\hat{\rho}_j^2}.$$

This gives us the distribution function,

$$F(\chi^2 - \Sigma\hat{\rho}) = \frac{1}{2\Gamma[\frac{1}{2}(n-m)]}\left[\tfrac{1}{2}(\chi^2 - \Sigma\hat{\rho}_j^2)\right]^{\frac{1}{2}(n-m)-1} e^{-\frac{1}{2}(\chi^2 - \Sigma\hat{\rho}_j^2)},$$

which is a chi-square distribution with $n - m$ degrees of freedom.

3. Application. As an example of the use of linear restrictions on chi-square we shall now examine the effect on the chi-square test of goodness of fit if the moments of a sample are not corrected for grouping errors in fitting a frequency curve.

The parameters of the fitted frequency distribution, $f(x)$, are determined from the equations,

$$N\int x^k f(x)\, dx = \Sigma\, x_j^k \theta_j, \qquad k = 0, 1, 2, \cdots, \tag{5}$$

where x_j is the mid-point of the j^{th} group and θ_j the corresponding observed frequency. Next a set of expected frequencies,

$$\hat{\theta}_j = \int_{\alpha_j}^{\alpha_{j+1}} Nf(x)\, dx, \qquad \alpha_j = (x_{j-1} + x_j)/2,$$

is determined by taking partial areas of the fitted frequency distribution. The expected frequency is used to transform the actual frequency into a statistic with mean zero and unit variance by the equation,

$$\chi_j = (\theta_j - \hat{\theta}_j)/\hat{\theta}^{\frac{1}{2}}.$$

Equations (5) may now be rearranged into the form of linear restrictions on the χ_j. Thus

$$\Sigma\, x_j^k \hat{\theta}_j^{\frac{1}{2}} \chi_j = \rho_k' \tag{6}$$

where the ρ_k' have the values,

$$\begin{aligned}\rho_k' &= \sum x_j^k \theta_j - \Sigma\, x_j^k \hat{\theta}_j \\ &= N\int x^k f(x)\, dx - \Sigma\, x_j^k \hat{\theta}_j \\ &\neq 0 \text{ in general}\end{aligned}$$

To make our example more specific, let us fit a normal distribution to a sample of 1000 items with mean zero and unit variance. Let the grouping be about the midpoints,

x_j: −3, −2, −1, 0, 1, 2, 3.

The expected frequencies in each group are

$\hat{\theta}_j$: 6, 61, 242, 382, 242, 61, 6.

The variance of these expected frequencies is 1.080 as contrasted with 1.000 for the sample. The linear restrictions, (6), now take the forms,

$$(7)\quad 2.4\chi_{-3} + 7.8\chi_{-2} + 15.6\chi_{-1} + 19.5\chi_0 + 15.6\chi_1 + 7.8\chi_2 + 2.4\chi_3 = 0$$

$$(8)\quad -7.2\chi_{-3} - 15.6\chi_{-2} - 15.6\chi_{-1} + 0 + 15.6\chi_1 + 15.6\chi_2 + 7.2\chi_3 = 0$$

$$(9)\quad 21.6\chi_{-3} + 31.2\chi_{-2} + 15.6\chi_{-1} + 0 + 15.6\chi_1 + 31.2\chi_2 + 21.6\chi_3 = -80.$$

Because of the symmetry of the normal distribution, restriction (8) is orthogonal to (7) and (9). Therefore the only orthogonalization necessary is to replace (9) by an equivalent restriction which is orthogonal to (7). This can be done by subtracting 1.080 times (7) from (9) which gives

$$(10)\quad 19.0\chi_{-3} + 22.8\chi_{-2} - 1.2\chi_{-1} - 21.1\chi_0 - 1.2\chi_1 + 22.8\chi_2 + 19.0\chi_3 = -80$$

If these restrictions are each divided by the square root of the sum of the squares of the coefficients of the χ_j, they will be the normal orthogonal set required by the development. The distances of these restrictive planes from the center of χ^2-sphere are

$$\rho_{(7)} = 0, \qquad \rho_{(8)} = 0, \qquad \rho_{(10)} = 1.7.$$

Thus if we test the goodness of fit of the normal distribution to this sample by calculating chi-square,

$$\chi^2 = \Sigma\, \chi_j^2 = \Sigma \frac{(\theta_j - \hat{\theta}_j)^2}{\hat{\theta}_j},$$

we should subtract from χ^2 a correction of

$$\Sigma\, \rho_k^2 = 2.8$$

before judging the significance. This correction adjusts for the effect of the grouping error on the chi-square test.

In this example, chi-square has four degrees of freedom so that an error of 2.8 is large enough to affect our judgment of its significance. It can be shown that the correction is proportional to the size of the sample. Therefore, if our sample had contained only 100 items, the fit obtained by ignoring grouping effects would be almost as good as the fit when the sample moments were corrected for grouping. On the other hand, if the sample had 10,000 items, it would be practically impossible to obtain a satisfactory fit without correcting for grouping errors.

4. Conclusion. The theory of the loss of degrees of freedom for chi-square when the underlying statistics are subject to linear restrictions does not require the restrictions to be homogeneous. For restrictions which are not homogeneous, a correction must be subtracted from chi-square equal to the square of the distance from the center of the sphere,

$$\chi^2 = \Sigma \chi_j^2 = 0$$

to the intersection of the restrictive planes. Non-homogeneous restrictions sometimes arise in practice because of the bias introduced by an approximation. An example is given from curve fitting.