Cochran's Q-test

Set Up. Each of c treatments is independently applied to r blocks (or subjects) and each outcome X_{ij} (where $i = 1, ..., r$ and $j = 1, ..., c$) is measured as a success $(X_{ij} = 1)$ or as a failure $(X_{ij} = 0)$.

	Treatments				
Blocks	1	2	...	c	row totals
1	X_{11}	X_{12}	...	X_{1c}	R_1
2	X_{21}	X_{22}	...	X_{2c}	R_2
$\ddots$	$\ddots$		$\ddots$		
r	X_{r1}	X_r	...	X_{rc}	R_r
C_j's	C_1	C_2	...	C_c	N

Note. delete all blocks containing all 0's and all blocks containg all 1's.

Hypotheses.

H_0: treatments are similarly effective $(p_{i1} = p_{i2} = ... = p_{ic}$, see below)
H_1: treatments differ in effectiveness

Let $X_{ij} = 1$ if treatment is effective and 0 otherwise. Let $p_{ij} = P(X_{ij} = 1)$ so that $X_{ij} \sim$ binomial(1, p_{ij}). Under H_0, $p_{i1} = p_{i2} = ... = p_{ic} \equiv p_i$ (say) for $i = 1,..., r$, where r is the number of blocks. Under H_0, each $X_{ij} \sim$ binomial(1, p_i) and upon applying Liapounov's central limit theorem (see below) when r is large, the column totals have an approximate normal distribution: $C_j = \sum_{i=1}^{r} X_{ij} \sim$ Normal. We then standardize the C_j, square the standardized result, and add the squares to obtain a test statistic that has an approximate χ^2 distribution.

In standardizing the C_j, we use the following estimates of $E(C_j)$ and Var(C_j). The mean $E(C_j)$ is estimated by the sample mean $\frac{\sum C_j}{c} = \frac{N}{c}$, where N denotes the total number of 1's. Since $\text{Var}(C_j) = \sum_{i=1}^{r} p_{ij}(1 - p_{ij})$, we can estimate it by $\sum \frac{R_i}{c}(1 - \frac{R_i}{c})\left(\frac{c}{c-1}\right)$ where $\frac{R_i}{c}$ is the natural estimate of p_i under H_0 and where $\frac{c}{c-1}$ is a correction factor that safeguards against type I errors. In estimating the mean and variance in this context, we lose a degree of freedom for the χ^2 distribution.

Test Statistic. The test statistic we've derived is

$$Q = \sum_{j=1}^{c} \left(\frac{C_j - \frac{N}{c}}{\sqrt{\sum \frac{c}{c-1} \frac{R_i}{c} \left(1 - \frac{R_i}{c}\right)}} \right)^2 = c(c-1) \frac{\sum_{j=1}^{c} \left(C_j - \frac{N}{c}\right)^2}{\sum R_i (c - R_i)} ,$$

or, equivalently,

$$Q = \frac{c(c-1)\sum_{j=1}^{c} C_j^2 - (c-1)N^2}{cN - \sum_{i=1}^{r} R_i^2}$$

which , under H_0, has an approximate χ^2 distribution with $c - 1$ degrees of freedom. A rule of thumb for a decent approximation is that we have at least $r = 4$ blocks and $rc \geq 24$.

Example. (from Daniel's text). Four methods (A, B, C, and D) of treating raw fabric to make it water repellent were tested for effectiveness on six types of fabric. A satisfactory result got a 1.

	Methods			
Fabric	A	B	C	D
I	1	1	0	0
II	1	1	0	1
III	1	0	0	0
IV	1	1	1	0
V	1	1	0	1
VI	1	1	0	1

Hypothesis. H_0 : The four treatments are equally effective.
H_1 : The four treatments are not equally effective.

Test statistic. $Q = \frac{4(3)[6^2+5^2+1^2+3^2]-3(15^2)}{4(15)-[2^2+3^2+1^2+3^2+3^2+3^2]} = 9.3158$

Decision rule. Reject H_0 if the p-value is sufficiently small.

***p*-value** $= P(\chi^2_3 \geq 9.3158) = .0254.$

Conclusion. We have sufficient statistical evidence (with p-value =.0254) to conclude that the effectiveness of at least two treatments differ.

Appendix

Liapounov Theorem (a version of the Central Limit Theorem).

Let X_1, X_2, ... be independent (not necessarily identically distributed) random variables with $E[X_n] = \mu_n$ and $\sigma^2 (X_n) = \sigma_n^2$. If $E(|X_i - \mu_i|^3) < \infty$ for i = 1,...,n and

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} E(|X_i - \mu_i|^3)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} = 0,$$

then

$$Z_n = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=}^{n} \sigma_i^2}}$$ is asymptotically distributed N(0,1).

(The proof involves the Cauchy-Schwarz inequality. See Edward Manoukian, *Mathematical Nonparametric Statistics*, 1986, Gordon & Breach.)