Four Frameworks for (2×2) Contingency Tables

i. Four Poisson streams with respect to a (2×2) cross-classification.
ii. Single multinomial for cross-classification of two binary responses.
iii. Two independent binomial distributions for a binary response.
iv. Hypergeometric distribution for binary response under randomized assignment of subjects to two equivalent treatments.

1. Four Poisson streams (e.g., number of occurrences of some rare event such as accidents, diseases, injuries, etc.).

Let $h = 1, 2$ and $i = 1, 2$ index two binary factors. Let n_{hi} denote the number of events for the $(hi)^{\text{th}}$ situation (e.g., sex $\times$ region).

Assume the $\{n_{hi}\}$ have independent Poisson distributions with $E\{n_{hi}\} = \mu_{hi}$.

The likelihood is

$$Pr[\{n_{hi}\}] = \prod_{h=1}^{2} \prod_{i=1}^{2} \{\mu_{hi}^{n_{hi}} e^{-\mu_{hi}} / n_{hi}!\}$$

Often the $\{\mu_{hi}\}$ have the form $\mu_{hi} = N_{hi}\lambda_{hi}$ where the $\{N_{hi}\}$ are total measures of exposure and the $\{\lambda_{hi}\}$ are rate parameters which are indicative of the rates at which events occur per unit of exposure.

Let $n = \sum_{h=1}^{2} \sum_{i=1}^{2} n_{hi}$. The total n has the Poisson distribution with $E\{n\} = \mu = \sum_{h=1}^{2} \sum_{i=1}^{2} \mu_{hi}$. This distribution follows from moment generating function methods. Thus, $Pr(n) = \mu^n e^{-\mu} / n!$.

The conditional distributions of the $\{n_{hi}\}$ given their total n is fixed is given by the following multnomial distribution

$$\begin{aligned} Pr[\{n_{hi}\} \mid n] &= \left\{ \prod_{h=1}^{2} \prod_{i=1}^{2} \mu_{hi}^{n_{hi}} e^{-\mu_{hi}} / n_{hi}! \right\} \Big/ \{\mu^n e^{-\mu} / n!\} \\ &= n! \prod_{h=1}^{2} \prod_{i=1}^{2} \left\{ \left(\frac{\mu_{hi}}{\mu} \right)^{n_{hi}} \Big/ n_{hi}! \right\} \\ &= n! \prod_{h=1}^{2} \prod_{i=1}^{2} \{\pi_{hi}^{n_{hi}} / n_{hi}!\} \\ &= \frac{n!}{n_{11}! n_{12}! n_{21}! n_{22}!} \pi_{11}^{n_{11}} \pi_{12}^{n_{12}} \pi_{21}^{n_{21}} \pi_{22}^{n_{22}} \end{aligned}$$

where $\pi_{hi} = (\mu_{hi} / \mu)$ and $(\pi_{11} + \pi_{12} + \pi_{21} + \pi_{22}) = 1$.

2. The single multinomial distribution for the cross-classification of two binary responses (e.g., a simple random sample from an essentially infinite population and the measurement of two binary characteristics such as sex and the presence or absence of a health condition). Let n denote sample size.

Let $y_{hiL} = \begin{cases} 1 & \text{if } L^{\text{th}} \text{ subject in sample has } (hi)^{\text{th}} \text{ response combination} \\ 0 & \text{if otherwise} \end{cases}$

Let $\pi_{hi} = E\{y_{hiL}\}$ denote the probability that a randomly selected subject has the $(hi)^{\text{th}}$ response combination or equivalently the proportion of the population with the $(hi)^{\text{th}}$ response combination.

$$Pr[\{y_{hiL}\}] = \prod_{h=1}^{2}\prod_{i=1}^{2}\pi_{hi}^{y_{hiL}} \text{ where} \sum_{h=1}^{2}\sum_{i=1}^{2}\pi_{hi} = 1.$$

This is the multivariate Bernoulli distribution.

Let $n_{hi} = \sum_{L=1}^{n} y_{hiL}$. Then n_{hi} is the number of subjects in the sample with the $(hi)^{\text{th}}$ response combination.

By moment generating function methods, the $\{n_{hi}\}$ can be shown to have the multinomial distribution

$$Pr[\{n_{hi}\}] = n!\prod_{h=1}^{2}\prod_{i=1}^{2}[\pi_{hi}^{n_{hi}}/n_{hi}!] \text{ where} \sum_{h=1}^{2}\sum_{i=1}^{2}\pi_{hi} = 1.$$

Thus, the multinomial distribution arises either as the conditional distribution of independent Poisson counts given their total or as the sum of independent multivariate Bernoulli distribution.

For the multinomial distribution, $E\{n_{hi}\} = n\pi_{hi}$.

$Var\{n_{hi}\} = n\pi_{hi}(1 - \pi_{hi})$.

$$Cov\{n_{hi}, n_{h'i'}\} = \sum_{L=1}^{n} Cov\{y_{hiL},\, y_{h'i'L}\}$$

$$= \sum_{L=1}^{n}[E\{y_{hiL}\, y_{h'i'L}\} - \pi_{hi}\pi_{h'i'}] = -n\pi_{hi}\pi_{h'i'}.$$

Let $p_{hi} = (n_{hi}/n)$ denote the sample proportion of subjects with the $(hi)^{\text{th}}$ response combination. Then

$$E\{p_{hi}\} = \pi_{hi},\ Var\{p_{hi}\} = \frac{\pi_{hi}(1-\pi_{hi})}{n},\ Cov\{p_{hi},\ p_{h'i'}\} = -\frac{\pi_{hi}\pi_{h'i'}}{n}.$$

Let $\boldsymbol{y}_{**L} = (y_{11L}, y_{12L}, y_{21L}, y_{22L})'$ denote the response vector for the L^{th} subject with respect to all four possible outcomes.

Each $\boldsymbol{y}'_{**L}$ has one of the following four forms: $(1,0,0,0),\ (0,1,0,0), (0,0,1,0),$ or $(0,0,0,1)$.

Let $\boldsymbol{n} = (n_{11}, n_{12}, n_{21}, n_{22})'$ denote the vector of observed frequencies for the respective response combinations.

$\boldsymbol{n} = \sum_{L=1}^{n} \boldsymbol{y}_{**L}$. Thus, $\boldsymbol{p} = (p_{11},\ p_{12}, p_{21}, p_{22})' = \frac{1}{n}\sum_{L=1}^{n} \boldsymbol{y}_{**L}$, i.e., the vector $\boldsymbol{p}$ of response proportions is a vector of sample means for indicator vectors $\{\boldsymbol{y}_{**L}\}$ for the n subjects.

$$\boldsymbol{E\{n\}} = \begin{bmatrix} n & \pi_{11} \\ n & \pi_{12} \\ n & \pi_{21} \\ n & \pi_{22} \end{bmatrix} = n\begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{21} \\ \pi_{22} \end{bmatrix} = n\boldsymbol{\pi}\text{; similarly, } E\{\boldsymbol{p}\} = \boldsymbol{\pi}$$

$$\boldsymbol{Var\{n\}} = \begin{bmatrix} n\pi_{11}(1-\pi_{11}) & -n\pi_{11}\pi_{12} & -n\pi_{11}\pi_{21} & -n\pi_{11}\pi_{22} \\ -n\pi_{11}\pi_{12} & n\pi_{12}(1-\pi_{12}) & -n\pi_{12}\pi_{21} & -n\pi_{12}\pi_{22} \\ -n\pi_{11}\pi_{21} & -n\pi_{12}\pi_{21} & n\pi_{21}(1-\pi_{21}) & -n\pi_{21}\pi_{22} \\ -n\pi_{11}\pi_{22} & -n\pi_{12}\pi_{22} & -n\pi_{21}\pi_{22} & n\pi_{22}(1-\pi_{22}) \end{bmatrix}$$

$$= n\begin{bmatrix} \pi_{11} & 0 & 0 & 0 \\ 0 & \pi_{12} & 0 & 0 \\ 0 & 0 & \pi_{21} & 0 \\ 0 & 0 & 0 & \pi_{22} \end{bmatrix} - n\begin{bmatrix} \pi_{11}^2 & \pi_{11}\pi_{12} & \pi_{11}\pi_{21} & \pi_{11}\pi_{22} \\ \pi_{12}\pi_{11} & \pi_{12}^2 & \pi_{12}\pi_{21} & \pi_{12}\pi_{22} \\ \pi_{21}\pi_{11} & \pi_{12}\pi_{21} & \pi_{21}^2 & \pi_{21}\pi_{22} \\ \pi_{22}\pi_{11} & \pi_{12}\pi_{22} & \pi_{21}\pi_{22} & \pi_{22}^2 \end{bmatrix}$$

$$= n\begin{bmatrix} \pi_{11} & 0 & 0 & 0 \\ 0 & \pi_{12} & 0 & 0 \\ 0 & 0 & \pi_{21} & 0 \\ 0 & 0 & 0 & \pi_{22} \end{bmatrix} - n\begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{21} \\ \pi_{22} \end{bmatrix}\begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{21} & \pi_{22} \end{bmatrix}$$
$$= n[\boldsymbol{D\pi} - \boldsymbol{\pi\pi'}]$$

where $\boldsymbol{D\pi}$ is a diagonal matrix with elements of $\boldsymbol{\pi}$ on diagonal. It follows that $\boldsymbol{Var\{p\}} = \frac{1}{n^2}\boldsymbol{Var\{n\}} = \frac{1}{n}[\boldsymbol{D\pi} - \boldsymbol{\pi\pi'}]$. Thus, $\boldsymbol{p}$ is an unbiased estimator for $\boldsymbol{\pi}$ because $E\{\boldsymbol{p}\} = \boldsymbol{\pi}$**.** Also, $\boldsymbol{p}$ is a consistent estimator for $\boldsymbol{\pi}$ because $\boldsymbol{Var\{p\}} = \frac{1}{n}[\boldsymbol{D\pi} - \boldsymbol{\pi\pi'}] \to 0$ as $n \to \infty$; i.e., $\boldsymbol{p} \to \boldsymbol{\pi}$ as $n \to \infty$.

Since $\boldsymbol{n} = (n_{11}, n_{12}, n_{21}, n_{22})'$ has the multinomial distribution with probability parameters $\boldsymbol{\pi} = (\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22})'$, the sum of its components $n_{1+} = (n_{11} + n_{12})$ has the binomial distribution with probability parameter $\pi_{1+} = (\pi_{11} + \pi_{12})$. This can be shown by moment generating function methods; more simply the response outcomes "11" or "12" can be viewed as "successes" and so $n_{1+} = (n_{11} + n_{12})$ becomes the observed number of successes in n trials for which the probability of success on each is $\pi_{1+} = (\pi_{11} + \pi_{12})$.

Thus, $Pr\{n_{1+}\} = \dfrac{n!}{n_{1+}!n_{2+}!}\pi_{1+}^{n_{1+}}\pi_{2+}^{n_{2+}}$.

If $\boldsymbol{n} = (n_{11}, n_{12}, n_{21}, n_{22})'$ has the multinomial distribution with probability parameters $\boldsymbol{\pi} = (\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22}),'$ then the conditional distribution of $\boldsymbol{n}$ given n_{1+} and hence also $n_{2+} = (n - n_{1+})$ has the following structure:

$$Pr[\{n_{hi}\}|n_{1+}, n_{2+}] = \frac{n!\pi_{11}^{n_{11}}\pi_{12}^{n_{12}}\pi_{21}^{n_{21}}\pi_{22}^{n_{22}}}{n_{11}!n_{12}!n_{21}!n_{22}!} \bigg/ \frac{n!\pi_{1+}^{n_{1+}}\pi_{2+}^{n_{2+}}}{n_{1+}!n_{2+}!}$$

$$= \left[\frac{n_{1+}!}{n_{11}!n_{12}!}\left(\frac{\pi_{11}}{\pi_{1+}}\right)^{n_{11}}\left(\frac{\pi_{12}}{\pi_{1+}}\right)^{n_{12}}\right]\left[\frac{n_{2+}!}{n_{21}!n_{22}!}\left(\frac{\pi_{21}}{\pi_{2+}}\right)^{n_{21}}\left(\frac{\pi_{22}}{\pi_{2+}}\right)^{n_{22}}\right]$$

$$= \left[\frac{n_{1+}!}{n_{11}!n_{12}!}\theta_{11}^{n_{11}}(1-\theta_{11})^{n_{12}}\right]\left[\frac{n_{2+}!}{n_{21}!n_{22}!}\ \theta_{21}^{n_{21}}(1-\theta_{21})^{n_{22}}\right]$$

where $\theta_{11} = (\pi_{11}/\pi_{1+})$ and $\theta_{21} = (\pi_{21}/\pi_{2+})$ are the conditional probabilities that $i = 1$ given respectively $h = 1$ and $h = 2$.

The structure of $Pr[\{n_{hi}\}|n_{1+},\ n_{2+}]$ is the same as the product of two binomial distributions, one involving n_{1+} trials with success probability θ_{11} and the other involving n_{2+} trials with success probability θ_{21}.

It is called the product binomial distribution.

3. The product binomial distribution of a single binary response for two independent samples (e.g., a stratified simple random sample from the two essentially infinite strata of a population (e.g., regions) and the measurement of one binary characteristic such as the presence or absence of a health condition).

Let n_{1+} and n_{2+} denote the sample sizes from the two strata indexed by $h = 1, 2$.

Let $u_{hiL} = \begin{cases} 1 & \text{if } L^{\text{th}} \text{ subject in sample from } h^{\text{th}} \text{ stratum has } i^{\text{th}} \text{ response} \\ 0 & \text{if otherwise} \end{cases}$

Let $\theta_{h1} = E\{u_{h1L}\}$ denote the probability that a randomly selected subject from the h^{th} stratum has the $i = 1$ response outcome or equivalently the proportion of the h^{th} stratum with the $i = 1$ response outcome.

Let $n_{h1} = \sum_{L=1}^{n_{h+}} u_{h1L}$ = the number of subjects in sample from h^{th} stratum with $i = 1$ response outcome.

The counts n_{11} and n_{21} have independent binomial distributions with probability parameters θ_{11} and θ_{21}. Thus, n_{11} and n_{21} jointly have the product binomial distribution

$$Pr[n_{11},\ n_{21}] = \left[\tfrac{n_{1+}!}{n_{11}!n_{12}!}\theta_{11}^{n_{11}}(1-\theta)^{n_{12}}\right]\left[\tfrac{n_{2+}!}{n_{21}!n_{22}!}\theta_{21}^{n_{21}}(1-\theta_{21})^{n_{22}}\right]$$

$$= Pr[\{n_{hi}\}] \text{ since } n_{12} = n_{1+} - n_{11},\ n_{22} = n_{2+} - n_{21}.$$

For the binomial distribution, $E\{n_{h1}\} = n_{h+}\theta_{h1}$ and $Var\{n_{h1}\} = n_{h+}\theta_{h1}(1-\theta_{h1})$. $Cov\{n_{11}, n_{21}\} = 0$ since stratified sampling involves independent selection for $h = 1, 2$.

Let $p_{h1} = (n_{h1}/n_{h+})$ denote sample proportion with $i = 1$ response outcome for h^{th} stratum.

$$E\{p_{h1}\} = \theta_{h1},\ Var\{p_{h1}\} = \tfrac{\theta_{h1}(1-\theta_{h1})}{n_{h+}},\ Cov\{p_{h1}, p_{h2}\} = 0.$$

4. If the probability parameters θ_{11} and θ_{21} are the same for the two strata, then $(n_{11} + n_{21}) = n_{+1}$ has the binomial distribution with their common value $\theta_{*1} = \theta_{11} = \theta_{21}$ as the probability parameter. This follows from the sum of two independent binomial random variables with the same probability parameter having the binomial distribution with that probability parameter.

Thus, $Pr[n_{+1}] = \frac{n!}{n_{+1}!n_{+2}!}\theta_{*1}^{n_{+1}}(1-\theta_{*1})^{n_{+2}}$.

The conditional distribution of n_{11} given n_{+1} is then

$$Pr[n_{11}|n_{+1},\, \theta_{11} = \theta_{21} = \theta_{*1}] = \frac{\frac{n_{1+}!}{n_{11}!n_{12}!}\theta_{*1}^{n_{11}}(1-\theta_{*1})^{n_{12}}\frac{n_{2+}!}{n_{21}!n_{22}!}\theta_{*1}^{n_{21}}(1-\theta_{*1})^{n_{22}}}{\frac{n!}{n_{+1}!n_{+2}!}\theta_{*1}^{n+1}(1-\theta_{*1})^{n+2}}$$

$$= \frac{n_{1+}!n_{2+}!n_{+1}!n_{+2}!}{n!n_{11}!n_{12}!n_{21}!n_{22}!} = \binom{n_{+1}}{n_{11}}\binom{n_{+2}}{n_{12}}\Big/\binom{n}{n_{1+}}$$

$$= Pr[\{n_{hi}\}|n_{1+}, n_{+1}, n, \theta_{11} = \theta_{21} = \theta_{*1}]$$

since $n_{12} = n_{1+} - n_{11}$, $n_{21} = n_{+1} - n_{11}$, and $n_{22} = n - n_{1+} - n_{+1} + n_{11}$.

The structure $Pr[n_{11}|n_{+1},\, \theta_{11} = \theta_{21} = \theta_{*1}]$ is the hypergeometric distribution. For this distribution, it can be verified that $E\{n_{11}\} = \{n_{1+}n_{+1}/n\} = m_{11}$ and that $Var\{n_{11}\} = \{n_{1+}n_{2+}n_{+1}n_{+2}/n^2(n-1)\} = v_{11}$.

For sufficiently large samples, n_{11} approximately has the normal distribution with expected value m_{11} and variance v_{11}. Accordingly, $Q = (n_{11} - m_{11})^2/v_{11}$ has the $\chi^2(\text{df} = 1)$ distribution.

The (observed-expected) count difference $(n_{11} - m_{11})$ can also be expressed as

$$(n_{11} - m_{11}) = \left(n_{11} - \frac{n_{1+}n_{+1}}{n}\right) = \left(\frac{n_{11}(n_{11}+n_{12}+n_{21}+n_{22}) - (n_{11}+n_{12})(n_{11}+n_{21})}{n}\right)$$

$$= \frac{(n_{11}n_{22} - n_{12}n_{21})}{n} = \frac{n_{1+}n_{2+}}{n}\left(\frac{n_{11}}{n_{1+}} - \frac{n_{21}}{n_{2+}}\right) = \frac{n_{1+}n_{2+}}{n}(p_{11} - p_{21});$$

and so large values of $(n_{11} - m_{11})^2$ correspond to substantial differences between p_{11} and p_{21}.

More generally, $E\{n_{hi}\} = \{n_{h+}n_{+i}/n\} = m_{hi}$

$$Var\{n_{hi}\} = \{n_{1+}n_{2+}n_{+1}n_{+2}/n^2(n-1)\} = v_{hi} = v_{11}$$

$$(n_{hi} - m_{hi}) = \pm(n_{11}n_{22} - n_{12}n_{21})/n = \pm(n_{11} - m_{11})$$

and so $Q = (n_{hi} - m_{hi})^2/v_{hi} = (n_{11} - m_{11})^2/v_{11}$; i.e., Q remains the same regardless of which cell of the (2×2) table is used for its construction.

If Q is sufficiently large that probability of larger values is less than some specified Type 1 error rate α, then either an unusual event has occurred or the framework leading to the hypergeometric distribution is contradicted; in the latter context, the condition $H_0 : \theta_{11} = \theta_{21}$ of equality of probability parameters for the two binomial distributions is said to be contradicted.

For large samples, the evaluation of H_0 can be based on the $\chi^2(\text{df} = 1)$ approximation for the distribution of Q. With small samples, outcomes $\{n_{hi}\}$ with probabilities equal to or less than the probability of the observed outcome are identified. If the sum of the probabilities of such events is less than α, then the hypothesis is contradicted by what is known as Fisher's exact test.

5. In a randomized clinical trial, a fixed population of n subjects is randomly assigned to two treatment groups.

Let n_{1+} denote the number assigned to Group A and let n_{2+} denote the number assigned to Group P.

If there is no difference between treatments for each subject in the sense that the potential responses to either treatment are identical for each subject, then n_{+1} can denote the number of subjects with favorable response regardless of treatment and n_{+2} can denote the number with unfavorable response regardless of treatment.

Subsequent to random assignment, n_{11} represents the number of subjects in Group A with favorable response.

Since the n_{1+} subjects in Group A are a simple random sample without replacement from the set of n subjects under study, the frequency n_{11} has the hypergeometric distribution

$$Pr[n_{11}|\ n_{1+},\ n_{+1}, n, \text{ no treatment difference}] = \frac{\binom{n_{+1}}{n_{11}}\binom{n_{+2}}{n_{12}}}{\binom{n}{n_{1+}}} = \frac{n_{1+}!n_{2+}!n_{+1}!n_{+2}!}{n!n_{11}!n_{12}!n_{21}!n_{22}!}.$$

If the sum of this probability and the probabilities of outcomes which are equally or less likely is less than or equal to the specified Type I error α, then the hypothesis of no treatment difference is contradicted by this application of Fisher's exact test.

Alternatively, for large samples, $Q = (n_{11} - m_{11})^2/v_{11}$ can be evaluated through its approximate $\chi^2(\text{df} = 1)$ distribution.

Thus, the hypergeometric distribution arises from either randomized assignment of subjects to treatment in the null hypothesis setting of no treatment difference or as the conditional distribution for binomial counts given their sum in the setting of no difference between their probability parameters.

The hypothesis of no treatment difference in the randomized clinical trial setting is thereby analagous to the hypothesis $\theta_{11} = \theta_{21}$ of equality of probability parameters of two binomial distributions which is analagous to the hypothesis

$$\left(\frac{\pi_{11}}{\pi_{1+}}\right) = \left(\frac{\pi_{21}}{\pi_{2+}}\right) \text{ or } \left(\frac{\pi_{11}}{\pi_{11}+\pi_{12}}\right) = \left(\frac{\pi_{21}}{\pi_{21}+\pi_{22}}\right)$$

$$\text{or } \Big(\pi_{11}(\pi_{21}+\pi_{22}) = \pi_{21}(\pi_{11}+\pi_{12})\Big) \text{ or } \Big(\pi_{11}\pi_{22} = \pi_{12}\pi_{21}\Big)$$

$$\text{or } \left(\frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = 1\right) \text{ or } \Big(\pi_{11}(1-\pi_{1+}-\pi_{+1}+\pi_{11}) = (\pi_{1+}-\pi_{11})(\pi_{+1}-\pi_{11})\Big)$$

$$\text{or } \pi_{11} = \pi_{1+}\pi_{+1} \text{ in which case } \pi_{12} = \pi_{1+}-\pi_{11} = \pi_{1+}\pi_{+2},$$

$$\pi_{21} = \pi_{+1}-\pi_{11} = \pi_{2+}\pi_{+1}, \text{ and } \pi_{22} = 1-\pi_{1+}-\pi_{+1}+\pi_{11} = \pi_{2+}\pi_{+2}$$

i.e., $\pi_{hi} = \pi_{h+}\pi_{+i}$ or the two binary responses for a single multinomial are independent of one another. This hypothesis is analagous to the hypothesis $\left(\frac{\mu_{11}\mu_{22}}{\mu_{12}\mu_{21}}\right) = 1$ or $\mu_{hi} = \exp(\mu + \xi_h + r_i)$ where μ is an overall mean, ξ_h is the effect for the h^{th} level of one factor for Poisson stream cross-classification and r_i is the effect for the i^{th} level of the other factor. Also, to avoid redundancy, it is specified that $\xi_1 + \xi_2 = 0,\ r_1 + r_2 = 0.$ The framework here for Poisson streams is said to be a log-linear model since

$$\begin{aligned} log_e\mu_{hi} &= \mu + \xi_h + r_i = \mu + \xi_1 + r_1 = log_e\mu_{11} \\ &= \mu + \xi_1 - r_1 = log_e\mu_{12} \\ &= \mu - \xi_1 + r_1 = log_e\mu_{21} \\ &= \mu - \xi_1 - r_1 = log_e\mu_{22} \end{aligned}$$

corresponds to a linear model. The linear model here specifies no interaction between two factors for Poisson streams.

Thus, the following hypotheses are equivalent for their corresponding models.

1. No interaction of two binary cross-classified factors in log-linear model for means of four independent Poisson streams.

2. Independence of two binary responses for single multinomial distribution resulting from simple random sample of essentially infinite population.

3. Homogeneity of probability parameters for two binomial distributions resulting from stratified simple random sample from two essentially infinite strata.

4. Equality of treatments in a randomized clinical trial where a binary response is observed.

For all of these settings

$$\begin{aligned} Q &= (n_{11} - m_{11})^2 / v_{11} \\ &= (p_{11} - p_{21})^2 \bigg/ \left\{ \frac{n_{+1} n_{+2}}{n_{1+} n_{2+} (n-1)} \right\} \\ &= (n_{11} n_{22} - n_{12} n_{21})^2 (n-1) / n_{1+} n_{2+} n_{+1} n_{+2} \\ &= \left(\frac{n-1}{n} \right) \sum_{h=1}^{2} \sum_{i=1}^{2} \frac{(n_{hi} - m_{hi})^2}{m_{hi}} = \left(\frac{n-1}{n} \right) Q_P \end{aligned}$$

approximately has the $\chi^2(\mathrm{df} = 1)$ distribution and this can be used for hypothesis evaluation. Alternatively, Fisher's exact test can be used on the basis of randomization in (4) and conditional distribution arguments in (1), (2), (3).

The statistic

$$Q_P = \sum_{h=1}^{2} \sum_{i=1}^{2} \frac{(n_{hi} - m_{hi})^2}{m_{hi}}$$

is known as the Pearson chi-square statistic for assessing independence of rows and columns of (2×2) table; it also approximately has $\chi^2(\mathrm{df} = 1)$ distribution.