Randomization Model Methods for One-Sample Repeated Measures

- For categorical response variables, the WLS approach is often inapplicable
  - sample size may be too small
  - number of time points may be too large
- For continuous response variables:
  - the normality assumption may not be valid
  - the unstructured multivariate approach often has low power
  - repeated measures ANOVA requires restrictive covariance assumptions
  - choice of alternative covariance structure may not be obvious
- An alternative methodology is based on the randomization model and the multiple hypergeometric distribution

Advantages

- Useful for assessing strength of association between a response and a repeated measures factor in a relatively assumption-free context

- Applies to categorical or continuous outcomes

- Applicable when sample sizes are too small to warrant the use of large-sample methods

  sample size requirements for asymptotic tests apply to across-strata totals, rather than to within-strata totals

- Easily accommodates missing data

  (if missing completely at random)

- Does not require random sampling of subjects from some underlying probabilistic framework

Disadvantages

- Provides hypothesis testing procedures only

  Estimation of parameters and construction of confidence intervals is not generally possible

- Not useful for modeling

  cannot assess influences of multiple factors

- Limited to one-sample problems

- The scope of inference is restricted to the actual subjects under study

  rather than to some broad population which the subjects might conceptually represent

- Tests may be insensitive to alternatives in which associations vary in direction across strata (subjects)

Randomization Model Methods for One-Sample Repeated Measures

- Based on the use of Cochran–Mantel–Haenszel statistics
  - Landis et al. (1978)
  - Landis et al. (1988)
  - Crowder and Hand (1990, Section 8.6)

- The methodology will be developed as follows:
  a. The hypergeometric distribution
  b. Large-sample tests of randomness for a single 2×2 table and for sets of 2×2 tables
  c. Repeated measures with a binary outcome
  d. The multiple hypergeometric distribution
  e. Large-sample tests of randomness for a single $r \times c$ table and for sets of $r \times c$ tables
  f. Repeated measures with general outcomes

The Hypergeometric Distribution

- Consider a population of n objects, of which $n_{.1}$ are of type 1 and $n - n_{.1}$ are of type 2
- Suppose that a sample of size $n_{1.}$ is selected from this population (without replacement)
- Let X denote the number of type 1 objects in the sample
- These data can be displayed in the following 2×2 table:

Sampled	Type 1	Type 2	Total
Yes	X	$n_{1.}$–X	$n_{1.}$
No	$n_{.1}$–X	n–$n_{1.}$–$n_{.1}$+X	n–$n_{1.}$
Total	$n_{.1}$	n–$n_{.1}$	n

- We write $X \sim H(n, n_{1.}, n_{.1})$

The Hypergeometric Distribution

- The distribution of $X \sim H(n, n_{1.}, n_{.1})$ is given by $h(x) = \Pr(X = x)$, where:

$$h(x) = \binom{n_{.1}}{x}\binom{n - n_{.1}}{n_{1.} - x} \Big/ \binom{n}{n_{1.}}$$

$$= \frac{\dfrac{n_{.1}!}{x!(n_{.1} - x)!}\dfrac{(n - n_{.1})!}{(n_{1.} - x)!(n - n_{.1} - n_{1.} + x)!}}{\dfrac{n!}{n_{1.}!(n - n_{1.})!}}$$

$$= \frac{n_{.1}!\,(n - n_{.1})!\,n_{1.}!\,(n - n_{1.})!}{n!\,x!(n_{.1} - x)!(n_{1.} - x)!(n - n_{.1} - n_{1.} + x)!}$$

for $\max(0, n_{1.} + n_{.1} - n) \leq x \leq \min(n_{1.}, n_{.1})$

- It can be shown that

$$\mathrm{E}(X) = \frac{n_{1.}n_{.1}}{n},$$

$$\mathrm{Var}(X) = \frac{n_{1.}(n - n_{1.})n_{.1}(n - n_{.1})}{n^2(n - 1)}$$

Test of Randomness for a 2×2 Table

- Consider a sample of n observations classified with respect to two dichotomous variables

- The resulting frequencies can be displayed in a 2×2 contingency table:

Row Variable	Column Variable		
	Level 1	Level 2	Total
Level 1	n_{11}	n_{12}	$n_{1.}$
Level 2	n_{21}	n_{22}	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	n

- If the row and column marginal totals are fixed (either by design or by conditioning), $n_{11} \sim H(n, n_{1.}, n_{.1})$

Test of Randomness for a 2×2 Table

- Under the null hypothesis of randomness,

$$h(n_{11}) = \frac{n_{1.}!\, n_{2.}!\, n_{.1}!\, n_{.2}!}{n!\, n_{11}!\, n_{12}!\, n_{21}!\, n_{22}!},$$

for $\max(0, n_{1.} + n_{.1} - n) \leq n_{11} \leq \min(n_{1.}, n_{.1})$

- Under the null hypothesis of randomness,

$$\mathrm{E}(n_{11}) = \frac{n_{1.} n_{.1}}{n}$$

$$\mathrm{Var}(n_{11}) = \frac{n_{1.} n_{2.} n_{.1} n_{.2}}{n^2 (n-1)}$$

- A large-sample test of randomness is based on the statistic

$$Q = \frac{\big(n_{11} - \mathrm{E}(n_{11})\big)^2}{\mathrm{Var}(n_{11})},$$

which has an asymptotic χ^2_1 distribution

Test of Randomness for s 2×2 Tables

- Consider a set of s independent 2×2 tables, with the counts in the hth table denoted by:

	Column Variable		
Row Variable	Level 1	Level 2	Total
Level 1	n_{h11}	n_{h12}	$n_{h1.}$
Level 2	n_{h21}	n_{h22}	$n_{h2.}$
Total	$n_{h.1}$	$n_{h.2}$	n_h

- We wish to test the null hypothesis

  H_0: no association between the row and column variable in any of the s tables

- If the row and column marginal totals in each table are fixed, the n_{h11} are independent hypergeometric random variables

  $n_{h11} \sim H(n_h, n_{h1.}, n_{h.1})$

Test of Randomness for s 2×2 Tables

- If H_0 is true,

$$\mathrm{E}(n_{h11}) = \frac{n_{h1.} n_{h.1}}{n_h}$$

$$\mathrm{Var}(n_{h11}) = \frac{n_{h1.} n_{h2.} n_{h.1} n_{h.2}}{n_h^2 (n_h - 1)}$$

- Now let $X = \sum_{h=1}^{s} n_{h11}$

$$\mathrm{E}(X) = \sum_{h=1}^{s} \mathrm{E}(n_{h11}) = \sum_{h=1}^{s} \frac{n_{h1.} n_{h.1}}{n_h}$$

$$\mathrm{Var}(X) = \sum_{h=1}^{s} \mathrm{Var}(n_{h11}) = \sum_{h=1}^{s} \frac{n_{h1.} n_{h2.} n_{h.1} n_{h.2}}{n_h^2 (n_h - 1)}$$

- H_0 can then be tested using the statistic

$$Q = \frac{(X - \mathrm{E}(X))^2}{\mathrm{Var}(X)},$$

which has an asymptotic null χ_1^2 distribution

Test of Randomness for s 2×2 Tables

- Commonly known as the Mantel-Haenszel test
- The asymptotic null distribution is valid when:
  - s is small, if the $\{n_h\}$ are large
  - s is large, even if the $\{n_h\}$ are small
- Q is large when $n_{h11} - \mathrm{E}(n_{h11})$ is consistently positive or consistently negative across strata
- If $n_{h11} - \mathrm{E}(n_{h11})$ is positive in some strata and negative in others, the MH test will have low power for detecting an overall association
- A continuity correction is sometimes used
  - $(|X - \mathrm{E}(X)| - 0.5)^2$ for the numerator of Q
  - Recommended only when all $n_h = 2$

Application to Repeated Measures

- Suppose that a dichotomous outcome is measured at $t = 2$ time points for each of n subjects

  e.g. y_{ij} takes on the values $+$ or $-$,

  for $i = 1, \ldots, n, \;\; j = 1, 2$

- The data from subject i can be displayed in a 2×2 contingency table:

	Response Category		
Time	$+$	$-$	Total
1	n_{i11}	n_{i12}	1
2	n_{i21}	n_{i22}	1
Total	$n_{i.1}$	$n_{i.2}$	2

- Note that in each table, two of the n_{ijk} values will be equal to 0 and two will be equal to 1

- In fact, there are only four possible tables

Application to Repeated Measures

Type of Table	No. of Subjects	$E(n_{i11})$	$\text{Var}(n_{i11})$
Response Time + − Total 1 1 0 1 2 1 0 1 Total 2 0 2	a	1	0
Response Time + − Total 1 1 0 1 2 0 1 1 Total 1 1 2	b	1/2	1/4
Response Time + − Total 1 0 1 1 2 1 0 1 Total 1 1 2	c	1/2	1/4
Response Time + − Total 1 0 1 1 2 0 1 1 Total 0 2 2	d	0	0

Application to Repeated Measures

$$
\begin{aligned}
X &= \sum_{i=1}^{n} n_{i11} \\
&= (a \times 1) + (b \times 1) + (c \times 0) + (d \times 0) \\
&= a + b
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_{i=1}^{n} \mathrm{E}(n_{i11}) \\
&= \Big(a \times 1\Big) + \Big(b \times \frac{1}{2}\Big) + \Big(c \times \frac{1}{2}\Big) + \Big(d \times 0\Big) \\
&= a + \frac{b+c}{2}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(X) &= \sum_{i=1}^{n} \mathrm{Var}(n_{i11}) \\
&= \Big(a \times 0\Big) + \Big(b \times \frac{1}{4}\Big) + \Big(c \times \frac{1}{4}\Big) + \Big(d \times 0\Big) \\
&= \frac{b+c}{4}
\end{aligned}
$$

Application to Repeated Measures

- Finally, we have

$$\begin{aligned} Q &= \frac{\big(X - \mathrm{E}(X)\big)^2}{\mathrm{Var}(X)} \\ &= \frac{\left(a + b - \left(a + \dfrac{b+c}{2}\right)\right)^2}{\dfrac{b+c}{4}} \\ &= \frac{(b-c)^2}{b+c} \end{aligned}$$

- In terms of the summary 2×2 table:

	Time 2		
Time 1	$+$	$-$	Total
$+$	a	b	$a+b$
$-$	c	d	$c+d$
Total	$a+c$	$b+d$	n

the test based on the statistic Q is equivalent to McNemar's test

Sample Size Considerations

- For the general case of s 2×2 tables, Mantel and Fleiss (1980) proposed a validity criterion for the Mantel-Haenszel statistic Q

- The minimum and maximum possible values of n_{h11} are:

  $$L_h = \max(0, n_{h11} - n_{h22}), \quad U_h = \min(n_{h1.}, n_{h.1})$$

- Provided that each of the two quantities

  $$\sum_{h=1}^{s} \mathrm{E}(n_{h11}) - \sum_{h=1}^{s} L_h, \quad \sum_{h=1}^{s} U_h - \sum_{h=1}^{s} \mathrm{E}(n_{h11})$$

  exceeds 5, the χ^2_1 distribution should adequately approximate the exact distribution of Q

- In the repeated measures setting, this requirement simplifies to $b + c \geq 10$

Example

- Although insulin pump therapy improves control of blood glucose levels in diabetic patients, side effects have been reported
- The following data on the occurrence of diabetic ketoacidosis (DKA) were obtained:

Occurrence of DKA		
Time Period 1 (Before Pump)	Time Period 2 (Pump Therapy)	No. of Patients
No	No	128
No	Yes	7
Yes	No	19
Yes	Yes	7
Total		161

Reference

Mecklenburg, R. S. et al. (1984). Acute complications associated with insulin pump therapy: report of experience with 161 patients. *JAMA* **252**, 3265–3269.

SAS Statements

```
data a;
input (dka1 dka2 count)
      ($char3. +1 $char3. 4.);
cards;
No  No  128
No  Yes   7
Yes No   19
Yes Yes   7
;
data b; set a;
keep id time dka discord;
retain id 0;
do i=1 to count;
id=id+1;
discord=(dka1 ne dka2);
time=1; dka=dka1;  output;
time=2; dka=dka2;  output;
end;
proc freq;
tables id*time*dka / noprint cmh;
title1 'All Data';
data c; set b; if discord=1;
proc freq;
tables id*time*dka / noprint cmh;
title1 'Discordant Pairs Only';
```

Multiple Hypergeometric Distribution

- Consider a population of n objects, of which $n_{.1}$ are of type 1, ..., $n_{.t}$ are of type t

- Suppose that s successive random samples of size $n_{1.}, \ldots, n_{s.}$ are selected from this population (without replacement)

- Let X_{ij} denote the number of elements of type j in sample i, for $i = 1, \ldots, s, j = 1, \ldots, t$

- The probability that the ith sample contains x_{ij} elements of type j is given by

$$f(\{x_{ij}\}) = \frac{\prod_{i=1}^{s} n_{i.}! \prod_{j=1}^{t} n_{.j}!}{n! \prod_{i=1}^{s} \prod_{j=1}^{t} x_{ij}!}$$

- $X = (X_{11}, \ldots, X_{st})' \sim H\big(n, \{n_{i.}\}, \{n_{.j}\}\big)$

Multiple Hypergeometric Distribution

- It can be shown that

$$\mathrm{E}(X_{ij}) = \frac{n_{i.}n_{.j}}{n}$$

$$\mathrm{Var}(X_{ij}) = \frac{n_{i.}(n-n_{i.})n_{.j}(n-n_{.j})}{n^2(n-1)}$$

$$\mathrm{Cov}(X_{ij}, X_{ij'}) = \frac{-n_{i.}(n-n_{i.})n_{.j}n_{.j'}}{n^2(n-1)}$$

$$\mathrm{Cov}(X_{ij}, X_{i'j}) = \frac{-n_{i.}n_{i'.}n_{.j}(n-n_{.j})}{n^2(n-1)}$$

$$\mathrm{Cov}(X_{ij}, X_{i'j'}) = \frac{n_{i.}n_{i'.}n_{.j}n_{.j'}}{n^2(n-1)}$$

- A general expression for the variances and covariances is

$$\mathrm{Cov}(X_{ij}, X_{i'j'}) = \frac{n_{i.}(\delta_{ii'}n-n_{i'.})n_{.j}(\delta_{jj'}n-n_{.j'})}{n^2(n-1)},$$

where $\delta_{ij} = 1$ if $i = j$, 0 otherwise

Test of Randomness for an $r \times c$ Table

- Consider a sample of N observations classified with respect to two categorical variables

- The resulting frequencies can be displayed in an $r \times c$ contingency table

Row Variable	Column Variable					
	1	$\cdots$	j	$\cdots$	c	Total
1	n_{11}	$\cdots$	n_{1j}	$\cdots$	n_{1c}	$n_{1.}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
i	n_{i1}	$\cdots$	n_{ij}	$\cdots$	n_{ic}	$n_{i.}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
r	n_{r1}	$\cdots$	n_{rj}	$\cdots$	n_{rc}	$n_{r.}$
Total	$n_{.1}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.c}$	N

- If the row and column marginal totals are fixed (either by design or by conditioning), $\{n_{ij}\} \sim H\big(N, \{n_{i.}\}, \{n_{.j}\}\big)$

Test of Randomness for an $r \times c$ Table

- Let $n = (n_{11}, \ldots, n_{1c}, \ldots, n_{r1}, \ldots, n_{rc})'$ denote the $rc \times 1$ vector of observed frequencies

- Let $p_{*.} = (p_{1.}, \ldots, p_{r.})'$ denote the $r \times 1$ vector of row marginal proportions, where $p_{i.} = n_{i.}/N$

- Let $p_{.*} = (p_{.1}, \ldots, p_{.c})'$ denote the $c \times 1$ vector of column marginal proportions, where $p_{.j} = n_{.j}/N$

- Let $m = \mathrm{E}(n)$, where

$$m = (m_{11}, \ldots, m_{1c}, \ldots, m_{r1}, \ldots, m_{rc})'$$

and

$$m_{ij} = \mathrm{E}(n_{ij}) = \frac{n_{i.} n_{.j}}{N} = N p_{i.} p_{.j}$$

- Using matrix notation, $\mathrm{E}(n) = N(p_{*.} \otimes p_{.*})$

Test of Randomness for an $r \times c$ Table

- Let Σ denote the $rc \times rc$ variance-covariance matrix of n

- The elements of Σ are given by

$$\begin{aligned}\text{Cov}(n_{ij}, n_{i'j'}) &= \frac{n_{i.}(\delta_{ii'}N - n_{i'.})n_{.j}(\delta_{jj'}N - n_{.j'})}{N^2(N-1)} \\ &= \frac{N^2}{N-1}p_{i.}(\delta_{ii'} - p_{i'.})p_{.j}(\delta_{jj'} - p_{.j'})\end{aligned}$$

  where $\delta_{ij} = 1$ if $i = j$, 0 otherwise

- Using matrix notation,

$$\Sigma = \frac{N^2}{N-1}(D_{p_{*.}} - p_{*.}p'_{*.}) \otimes (D_{p_{.*}} - p_{.*}p'_{.*})$$

  where $D_{p_{*.}}$ and $D_{p_{.*}}$ are diagonal matrices with the elements of $p_{*.}$ and $p_{.*}$ on the main diagonal

Test of Randomness for an $r \times c$ Table

- The asymptotic distribution of $N^{-1/2}(n - m)$ is $N_{rc}\left(0, \frac{1}{N}\Sigma\right)$

- If the sample size N is large, $n \approx N_{rc}(m, \Sigma)$

- Let $A = (I_{r-1}, 0_{r-1}) \otimes (I_{c-1}, 0_{c-1})$

  I_u is the $u \times u$ identity matrix

  0_u is a $u \times 1$ vector of 0's

  A is a $(r-1)(c-1) \times rc$ matrix

- Let $G = A(n - m)$ denote the $(r-1)(c-1) \times 1$ vector of differences between the observed and expected frequencies (under the null hypothesis of randomness)

  The linear transformation matrix A eliminates the last row and last column

Test of Randomness for an $r \times c$ Table

- Under the null hypothesis of randomness,

$$\mathrm{E}(G) = 0_{(r-1)(c-1)}$$

$$\mathrm{Var}(G) = A\Sigma A'$$

- Since $G \approx N_{(r-1)(c-1)}(0, A\Sigma A')$ under H_0,

$$Q = G'(A\Sigma A')^{-1}G$$

  is the large-sample quadratic form statistic for testing H_0

- If H_0 is true, $Q \approx \chi^2_{(r-1)(c-1)}$

- It can be shown that

$$Q = \frac{N-1}{N}X^2,$$

  where X^2 is the Pearson chi-square statistic

Test of Randomness for s $r \times c$ Tables

- Consider a set of s independent $r \times c$ tables, with the counts in the hth table denoted by:

Row Variable	Column Variable			
	1	$\cdots$	c	Total
1	n_{h11}	$\cdots$	n_{h1c}	$n_{h1.}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
r	n_{hr1}	$\cdots$	n_{hrc}	$n_{hr.}$
Total	$n_{h.1}$	$\cdots$	$n_{h.c}$	N_h

- We wish to test the null hypothesis

  H_0: no association between the row and column variable in any of the s tables

- If the row and column marginals in each table are fixed, $n_h = (n_{h11}, \ldots, n_{hrc})'$ are independent multiple hypergeometric random variables

$$n_h \sim H\big(N_h, \{n_{hi.}\}, \{n_{h.j}\}\big)$$

Test of Randomness for s $r \times c$ Tables

- If H_0 is true, $\mathrm{E}(n_{hij}) = (n_{hi.}n_{h.j})/N_h$ and $\mathrm{Cov}(n_{hij}, n_{hi'j'}) =$

$$\frac{n_{hi.}(\delta_{ii'}N_h - n_{hi'.})n_{h.j}(\delta_{jj'}N_h - n_{h.j'})}{N_h^2(N_h - 1)}$$

- Let $p_{h*.} = (p_{h1.}, \ldots, p_{hr.})'$ denote the $r \times 1$ vector of row marginal proportions in the hth table, where $p_{hi.} = n_{hi.}/N_h$, for $i = 1, \ldots, r$

- Let $p_{h.*} = (p_{h.1}, \ldots, p_{h.c})'$ denote the $c \times 1$ vector of column marginal proportions in the hth table, where $p_{h.j} = n_{h.j}/N_h$, for $j = 1, \ldots, c$

- Using matrix notation,

$$m_h = \mathrm{E}(n_h) = N_h(p_{h*.} \otimes p_{h.*})$$

$$\Sigma_h = \frac{N_h^2}{N_h - 1}(D_{p_{h*.}} - p_{h*.}p_{h*.}') \otimes (D_{p_{h.*}} - p_{h.*}p_{h.*}')$$

Test of Randomness for s $r \times c$ Tables

- Let $A = (I_{r-1}, 0_{r-1}) \otimes (I_{c-1}, 0_{c-1})$

  I_u is the $u \times u$ identity matrix

  0_u is a $u \times 1$ vector of 0's

  A is a $(r-1)(c-1) \times rc$ matrix

- Let $G_h = A(n_h - m_h)$ denote the $(r-1)(c-1) \times 1$ vector of differences between the observed and expected frequencies (under the null hypothesis of randomness) in the hth table

- Let $G = \sum_{h=1}^{s} G_h$

- Since the s tables are independent,

$$\mathrm{E}(G) = \sum_{h=1}^{s} \mathrm{E}(G_h) = 0_{(r-1)(c-1)},$$

$$\mathrm{Var}(G) = V_G = \sum_{h=1}^{s} \mathrm{Var}(G_h) = \sum_{h=1}^{s} A\Sigma_h A'$$

CMH General Association Statistic

- Since $G \approx N_{(r-1)(c-1)}(0, V_G)$ under H_0, the large-sample quadratic form statistic for testing H_0 is $Q_G = G' V_G^{-1} G$

  the Cochran/Mantel-Haenszel/Birch statistic

- If H_0 is true, $Q_G \approx \chi^2_{(r-1)(c-1)}$

- The asymptotic distribution of Q_G is linked to the total sample size $N = \sum_{h=1}^{s} N_h$, rather than to the stratum-specific sample sizes

- Q_G can be used when the row and column variables are nominal

  The null hypothesis is tested in terms of $(r-1)(c-1)$ linearly independent functions of the observed counts

CMH General Association Statistic

- If the CMH statistic Q_G is significant, then there is an association between the row and column variables in at least one of the s strata

- However, the power of Q_G is directed towards average partial association alternatives

  If certain observed frequencies consistently exceed (or are exceeded by) their corresponding expected frequencies, then these quantities reinforce one another when combined across strata

- Q_G has low power for detecting associations which are not consistent across strata

- If $r = c = 2$, Q_G is the Mantel-Haenszel test

- If $s = 1$, then $Q_G = (1 - 1/N)X^2$

CMH Mean Score Statistic

- Consider a set of s independent $r \times c$ tables, with the counts in the hth table denoted by:

Row Variable	Column Variable 1	$\cdots$	c	Total
1	n_{h11}	$\cdots$	n_{h1c}	$n_{h1.}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
r	n_{hr1}	$\cdots$	n_{hrc}	$n_{hr.}$
Total	$n_{h.1}$	$\cdots$	$n_{h.c}$	N_h

- Suppose that the column variable is ordinal and that appropriate scores $b_{h1}, \ldots, b_{hc}$ can be assigned to the levels

- In this case, we may wish to test

  H_0: no association between the row and column variable in any of the s tables

  versus the alternative that the r mean scores differ, on average, across tables

CMH Mean Score Statistic

- Under H_0, and conditional on the row and column marginals in each table, $n_h = (n_{h11}, \ldots, n_{hrc})'$ are independent multiple hypergeometric random variables

$$n_h \sim H\big(N_h, \{n_{hi.}\}, \{n_{h.j}\}\big)$$

- If H_0 is true, $m_h = \mathrm{E}(n_h) = N_h(p_{h*.} \otimes p_{h.*})$ and $\Sigma_h = \mathrm{Var}(n_h) =$

$$\frac{N_h^2}{N_h - 1}(D_{p_{h*.}} - p_{h*.}p_{h*.}') \otimes (D_{p_{h.*}} - p_{h.*}p_{h.*}')$$

where

$$p_{h*.} = (p_{h1.}, \ldots, p_{hr.})', \text{ with } p_{hi.} = n_{hi.}/N_h$$

$$p_{h.*} = (p_{h.1}, \ldots, p_{h.c})', \text{ with } p_{h.j} = n_{h.j}/N_h$$

and $D_{p_{h*.}}$ and $D_{p_{h.*}}$ are diagonal matrices with the elements of $p_{h*.}$ and $p_{h.*}$ on the main diagonal

CMH Mean Score Statistic

- Let $A_h = (I_{r-1}, 0_{r-1}) \otimes (b_{h1}, \ldots, b_{hc})$

  I_u is the $u \times u$ identity matrix

  0_u is a $u \times 1$ vector of 0's

  A_h is a $(r-1) \times rc$ matrix

- Let $M_h = A_h(n_h - m_h)$ denote the $(r-1) \times 1$ vector of differences between the observed and expected mean scores (under the null hypothesis of randomness) in the hth table

- Let $M = \sum_{h=1}^{s} M_h$

- Since the s tables are independent,

$$\mathrm{E}(M) = \sum_{h=1}^{s} \mathrm{E}(M_h) = 0_{(r-1)},$$

$$\mathrm{Var}(M) = V_M = \sum_{h=1}^{s} \mathrm{Var}(M_h) = \sum_{h=1}^{s} A_h \Sigma_h A_h'$$

CMH Mean Score Statistic

- Since $M \approx N_{(r-1)}(0, V_M)$ under H_0, the large-sample quadratic form statistic for testing H_0 is $Q_M = M' V_M^{-1} M$

- If H_0 is true, $Q_M \approx \chi^2_{(r-1)}$

- The null hypothesis is tested in terms of $(r-1)$ linearly independent functions of the observed mean scores

- Q_M is directed at location-shift alternatives

  the extent to which the mean scores in certain rows consistently exceed (or are exceeded by) the mean scores in other rows

- If $s = 1$ and rank scores are used, Q_M is equivalent to the Kruskal-Wallis test

CMH Correlation Statistic

- Consider a set of s independent $r \times c$ tables, with the counts in the hth table denoted by:

Row Variable	Column Variable 1	$\cdots$	c	Total
1	n_{h11}	$\cdots$	n_{h1c}	$n_{h1.}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
r	n_{hr1}	$\cdots$	n_{hrc}	$n_{hr.}$
Total	$n_{h.1}$	$\cdots$	$n_{h.c}$	N_h

- Suppose that the row and column variables are both ordinal

$$\text{row scores: } a_{h1}, \ldots, a_{hr}$$

$$\text{column scores: } b_{h1}, \ldots, b_{hc}$$

- In this case, we may wish to test H_0 versus the alternative that there is a consistent positive (or negative) association between the row scores and the column scores, across tables

CMH Correlation Statistic

- Let $A_h = (a_{h1}, \ldots, a_{hr}) \otimes (b_{h1}, \ldots, b_{hc})$

  A_h is a row vector with rc components

- Let $C_h = A_h(n_h - m_h)$ denote the difference between the observed and expected association scores (under the null hypothesis of randomness) in the hth table

- Let $C = \sum_{h=1}^{s} C_h$

- Since the s tables are independent,

$$\mathrm{E}(C) = \sum_{h=1}^{s} \mathrm{E}(C_h) = 0,$$

$$\mathrm{Var}(C) = V_C = \sum_{h=1}^{s} \mathrm{Var}(C_h) = \sum_{h=1}^{s} A_h \Sigma_h A_h'$$

CMH Correlation Statistic

- Since $C \approx N(0, V_C)$ under H_0, the large-sample quadratic form statistic for testing H_0 is $Q_C = C^2/V_C$

- If H_0 is true, $Q_C \approx \chi^2_1$

- Q_C is directed at correlation alternatives

  the extent to which there is a consistent positive (or negative) linear association between the row and column scores

- If $s = 1$, then $Q_C = (N-1)r^2$, where r is the Pearson correlation coefficient between the row and column scores

Summary of CMH Statistics

- For s independent $r \times c$ tables, there are three CMH statistics:

Alternative Hypothesis	Stat.	df	Variable Type: Row	Variable Type: Column
General assoc.	Q_G	$(r\text{-}1)(c\text{-}1)$	nominal	nominal
Mean score differences	Q_M	$r-1$	nominal	ordinal
Linear assoc.	Q_C	1	ordinal	ordinal

- In repeated measures applications,

  Q_G tests marginal homogeneity across time

  Q_M tests equality of means across time

  Q_C tests for linear association between the response and time

Application to Repeated Measures

- Suppose that a categorical variable with c possible outcomes is measured at t time points for each of n subjects

  e.g. y_{ij} takes on the values $1, \ldots, c$,
  for $i = 1, \ldots, n$, $j = 1, \ldots, t$

- We wish to test if the marginal distribution of the response is the same at each of the t time points

- Define the indicator variables

$$n_{ijk} = \begin{cases} 1, & \text{if subject } i \text{ is classified in} \\ & \text{response category } k \text{ at time } j \\ 0, & \text{otherwise} \end{cases}$$

  for $i = 1, \ldots, n$, $j = 1, \ldots, t$, $k = 1, \ldots, c$

Application to Repeated Measures

- The data from subject i can be displayed in a $t \times c$ contingency table:

	Response Category			
Time	1	$\cdots$	c	Total
1	n_{i11}	$\cdots$	n_{i1c}	1
$\vdots$	$\vdots$		$\vdots$	$\vdots$
t	n_{it1}	$\cdots$	n_{itc}	1
Total	$n_{i.1}$	$\cdots$	$n_{i.c}$	t

- In each row of the table, one of the n_{ijk} values will be equal to 1 and the remaining n_{ijk} values will be equal to 0

- The column marginal total $n_{i.k}$ is the number of times that subject i was classified in response category k

  $0 \leq n_{i.k} \leq t$

Application to Repeated Measures

- Under the assumption that the column marginal totals $\{n_{i.k}\}$ are fixed, the null hypothesis of "no partial association" between the row dimension (time) and the column dimension (response) can be tested using Q_G

- In this context, there are n strata, one for each subject

- The "no partial association" hypothesis is the same as the "interchangeability" hypothesis of Madansky (1963)

- This null hypothesis implies marginal homogeneity in the distribution of the response across the t time points

Application to Repeated Measures

- Although the data in each table are sparse (all counts will be 0 or 1), the asymptotic distribution is linked to the total sample size $N = \sum_{h=1}^{s} N_h$

- For repeated measurement designs, the CMH statistic Q_G is equivalent to:

c	t	Method
2	2	McNemar's test
2	> 2	Cochrans's Q test
> 2	> 2	Birch's Lagrange multiplier test Madansky's interchangeability test

Score Options in SAS

Given the observed counts:

Row Variable	Column Variable 1	$\cdots$	c	Total
1	n_{h11}	$\cdots$	n_{h1c}	$n_{h1.}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
r	n_{hr1}	$\cdots$	n_{hrc}	$n_{hr.}$
Total	$n_{h.1}$	$\cdots$	$n_{h.c}$	N_h

SAS has four options for defining the row scores a_{hi} and the column scores b_{hj}

1. SCORES=TABLE (the default)

   if the row (column) variable is numeric, a_{hi} (b_{hj}) is the observed level for category i (j)

   if the row (column) variable is character, $a_{hi} = 1, 2, \ldots, r$ ($b_{hj} = 1, 2, \ldots, c$)

Score Options in SAS

2. SCORES=RANK

$$a_{hi} = R_{ahi} = \sum_{k=1}^{i-1} n_{hk.} + \frac{n_{hi.} + 1}{2}$$

$$b_{hj} = R_{bhj} = \sum_{k=1}^{j-1} n_{h.k} + \frac{n_{h.j} + 1}{2}$$

- These are the standard rank scores using midranks for tied observations
- If $s = 1$ and $r = 2$, Q_M and Q_C are the Mann-Whitney-Wilcoxon test
- If $s = 1$ and $r > 2$, Q_M is the Kruskal-Wallis test
- If $s > 1$ and $n_{hi.} = 1$, Q_M is Friedman's chi-square test

Score Options in SAS

3. SCORES=RIDIT

   $a_{hi} = R_{ahi}/N_h, \quad b_{hj} = R_{bhj}/N_h$

   This definition differs from the ridit scores

   $a_{hi} = (R_{ahi} - .5)/N_h, \quad b_{hj} = (R_{bhj} - .5)/N_h$

   defined by other authors

4. SCORES=MODRIDIT

$$a_{hi} = \frac{2\sum_{k=1}^{i} n_{hk.} - n_{hi.} + 1}{2(N_h + 1)} = \frac{R_{ahi}}{N_h + 1}$$

$$b_{hj} = \frac{2\sum_{k=1}^{j} n_{h.k} - n_{h.j} + 1}{2(N_h + 1)} = \frac{R_{bhj}}{N_h + 1}$$

   also known as standardized midrank scores

   yields van Elteren's (1960) test for combining Wilcoxon rank sum tests across a set of strata

Row and Column Scores in Repeated Measures Applications

- $t \times c$ contingency table for subject h:

	Response Category			
Time	1	$\cdots$	c	Total
1	n_{h11}	$\cdots$	n_{h1c}	1
$\vdots$	$\vdots$		$\vdots$	$\vdots$
t	n_{ht1}	$\cdots$	n_{htc}	1
Total	$n_{h.1}$	$\cdots$	$n_{h.c}$	t

- In this case, the scores are given by:

Scores	a_{hi}	b_{hj}
Table	i	j
Rank	i	R_{bhj}
Ridit	i/t	R_{bhj}/t
Modridit	$i/(t+1)$	$R_{bhj}/(t+1)$

- If there are no missing data, the results from rank, ridit, and modridit scores will be identical

Example

- 46 subjects were treated with each of three drugs (A, B, and C)
- The response to each drug was recorded as favorable (F) or unfavorable (U)
- The data from the ith subject can be displayed in a 3×2 contingency table:

	Response		
Drug	F	U	Total
A	n_{i11}	n_{i12}	1
B	n_{i21}	n_{i22}	1
C	n_{i31}	n_{i32}	1
Total	$n_{i.1}$	$n_{i.2}$	3

- The CMH statistic Q_G can be used to test

  H_0: for each subject, the total number of favorable responses ($n_{i.1}$) is distributed at random with respect to the three drugs

SAS Statements

```
data a;
input subject a $ b $ c $;
cards;
 1 F F U
   ...
46 U F U
 ;
data b; set a;
keep subject drug response;
drug='A';  response=a;  output;
drug='B';  response=b;  output;
drug='C';  response=c;  output;
proc freq;
tables subject*drug*response
       / noprint cmh;
```

- The results from this analysis are:

$$Q_G = 8.471, \quad \text{df} = 2, \quad p = .014$$

- We conclude that the response profiles of the three drugs are different

Example

- A study of the efficacy of steam inhalation in the treatment of common cold symptoms
- Eligible subjects had colds of recent onset (symptoms of nasal drainage, nasal congestion, and sneezing for 3 days or less)
- 32 patients were given two 20-minute steam inhalation treatments
- Severity of nasal drainage was self-assessed for four days

  0=no symptoms 2=moderate symptoms
  1=mild symptoms 3=severe symptoms

- Does symptom severity improve following treatment?

Reference

Macknin, M. L. et al. (1990). Effect of inhaling heated vapor on symptoms of the common cold. *JAMA* **264**, 989–991.

Severity of Nasal Drainage

ID	Day 1	Day 2	Day 3	Day 4
1	1	1	2	2
2	0	0	0	0
3	1	1	1	1
4	1	1	1	1
5	0	2	2	0
6	2	0	0	0
7	2	2	1	2
8	1	1	1	0
9	3	2	1	1
10	2	2	2	3
11	1	0	1	1
12	2	3	2	2
13	1	3	2	1
14	2	1	1	1
16	2	3	3	3
17	2	1	1	1
18	1	1	1	1
20	2	2	2	2
21	3	1	1	1
22	1	1	2	1
23	2	1	1	2
24	2	2	2	2
25	1	1	1	1
26	2	2	3	1
27	2	0	0	0
28	1	1	1	1
29	0	1	1	0
30	1	1	1	1
31	1	1	1	0
32	3	3	3	3

Analysis Options

- Normal-theory methods

  the response is not normally-distributed

- Weighted least squares approach

  Since there are $c^t = 4^4 = 256$ potential response profiles, the sample size is too small

- Randomization model methods

  Q_G with 9 df will have low power

  Since the response is ordinal, mean symptom scores across the four days can be compared using Q_M with 3 df

  Q_C can be used to test if there is a significant association between time and response

SAS Statements

- Compute Q_M and Q_C using the scores 1–4 for the row variable (time) and both the actual symptom scores (0–3) and rank scores for the column variable (drainage severity)

```
data a;
input id d1-d4;
cards;
 1 1 1 2 2
    ...
32 3 3 3 3
;
data b; set a;
day=1; drain=d1; output;
day=2; drain=d2; output;
day=3; drain=d3; output;
day=4; drain=d4; output;
proc freq;
tables id*day*drain
  / cmh noprint;
tables id*day*drain
  / cmh noprint scores=rank;
```

Accommodation of Missing Data
Drug Response Data from 46 Subjects

- The observed responses from subject 1 were:

  Drug A: F, Drug B: F, Drug C: U

- Now suppose that the drug B response was missing

- One approach would be to exclude this subject from the analysis

- In this case,

$$G = \sum_{h=2}^{46} G_h = \begin{pmatrix} 3.667 \\ 3.667 \end{pmatrix}$$

$$\mathrm{Var}(G) = \sum_{h=2}^{46} \mathrm{Var}(G_h) = \begin{pmatrix} 7.333 & -3.667 \\ -3.667 & 7.333 \end{pmatrix}$$

and $Q_G = G' \big(\mathrm{Var}(G)\big)^{-1} G = 7.333$

Accommodation of Missing Data

- However, the exclusion of subject 1 does not allow us to use the information that the response to Drug A (C) was favorable (unfavorable)
- Alternatively, the data from subject 1 can be displayed as follows:

	Response		
Drug	F	U	Total
A	1	0	1
B	0	0	0
C	0	1	1
Total	1	1	2

- In this case,

$$n_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad m_1 = \begin{pmatrix} .5 \\ .5 \\ .0 \\ .0 \\ .5 \\ .5 \end{pmatrix}$$

Accommodation of Missing Data

- The variance-covariance matrix of n_1 is

$$\Sigma_1 = \begin{pmatrix} .25 & -.25 & .00 & .00 & -.25 & .25 \\ -.25 & .25 & .00 & .00 & .25 & -.25 \\ .00 & .00 & .00 & .00 & .00 & .00 \\ .00 & .00 & .00 & .00 & .00 & .00 \\ -.25 & .25 & .00 & .00 & .25 & -.25 \\ .25 & -.25 & .00 & .00 & -.25 & .25 \end{pmatrix}$$

- The components of Q_G from subject 1 are

$$\begin{aligned} G_1 &= A(n_1 - m_1) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} .5 \\ -.5 \\ .0 \\ .0 \\ -.5 \\ .5 \end{pmatrix} \\ &= \begin{pmatrix} .5 \\ 0 \end{pmatrix} \end{aligned}$$

$$\mathrm{Var}(G_1) = A\Sigma_1 A' = \begin{pmatrix} .25 & 0 \\ 0 & 0 \end{pmatrix}$$

Accommodation of Missing Data

- Using the partial data from subject 1,

$$G = G_1 + \sum_{h=2}^{46} G_h$$

$$= \begin{pmatrix} .5 \\ 0 \end{pmatrix} + \begin{pmatrix} 3.667 \\ 3.667 \end{pmatrix}$$

$$= \begin{pmatrix} 4.167 \\ 3.667 \end{pmatrix}$$

$$\mathrm{Var}(G) = \mathrm{Var}(G_1) + \sum_{h=2}^{46} \mathrm{Var}(G_h)$$

$$= \begin{pmatrix} .25 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 7.333 & -3.667 \\ -3.667 & 7.333 \end{pmatrix}$$

$$= \begin{pmatrix} 7.583 & -3.667 \\ -3.667 & 7.333 \end{pmatrix}$$

- Thus, $Q_G = G' \big(\mathrm{Var}(G)\big)^{-1} G = 8.094$

Example

- The Muscatine Coronary Risk Factor Study
  - A longitudinal study of coronary risk factors in school children
- Dichotomous response (obese, not obese) obtained at three cross-sectional surveys
- Results from a cohort of 522 males who were 7–9 years old in 1977 are summarized below:

	All Data		Complete Cases	
Year	n	% Obese	n	% Obese
1977	356	18.8	225	19.6
1979	375	20.5	225	19.1
1981	380	23.7	225	23.1

- Is the marginal probability of obesity the same at each of the three years?

SAS Statements

```
data a;
input o77 o79 o81 count;
cards;
1 1 1  20
  ...
3 3 2  55
;
data b; set a;
keep id year obese complete; retain id 0;
if count>0 then do;
complete=(o77 ne 3)&(o79 ne 3)&(o81 ne 3);
if o77=3 then o77=.;
if o79=3 then o79=.;
if o81=3 then o81=.;
do i=1 to count; id=id+1;
year=77; obese=o77;  output;
year=79; obese=o79;  output;
year=81; obese=o81;  output;
end; end;
proc freq;
tables id*year*obese / noprint cmh;
data c; set b; if complete=1;
proc freq;
tables id*year*obese / noprint cmh;
```

Accommodation of Missing Data
Efficacy of Steam Inhalation

- The previous analysis excluded two subjects:

ID	Day 1	Day 2	Day 3	Day 4
15		3	3	2
19	3		1	0

- Both subjects' data support the hypothesis that symptoms improve over time and can be included in the computation of Q_M and Q_C

- First, the mean score statistic for the complete cases is computed as follows (using the actual symptom scores 0–3):

$$A_h = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 & 2 & 3 \end{pmatrix}$$

$$M_h = A_h(n_h - m_h)$$

for $h = 1, \ldots, 30$

Accommodation of Missing Data

- For the complete cases,

$$M = \sum_{h=1}^{30} M_h = \begin{pmatrix} 4.5 \\ 0.5 \\ 0.5 \end{pmatrix}$$

$$V_M = \sum_{h=1}^{30} A_h \Sigma_h A_h' = \begin{pmatrix} 7.750 & -2.583 & -2.583 \\ & 7.750 & -2.583 \\ & & 7.750 \end{pmatrix}$$

$$Q_M = M' V_M^{-1} M = 4.935 \qquad (\text{df}=3,\ p=0.177)$$

- The observed tables for subjects 15 and 19 are:

Subject 15

	Response				
Day	0	1	2	3	Sum
1	0	0	0	0	0
2	0	0	0	1	1
3	0	0	0	1	1
4	0	0	1	0	1
Sum	0	0	1	2	3

Subject 19

	Response				
Day	0	1	2	3	Sum
1	0	0	0	1	1
2	0	0	0	0	0
3	0	1	0	0	1
4	1	0	0	0	1
Sum	1	1	0	1	3

Accommodation of Missing Data

- The corresponding expected frequencies are:

Subject 15					
	Response				
Day	0	1	2	3	Sum
1	0	0	0	0	0
2	0	0	$\frac{1}{3}$	$\frac{2}{3}$	1
3	0	0	$\frac{1}{3}$	$\frac{2}{3}$	1
4	0	0	$\frac{1}{3}$	$\frac{2}{3}$	1
Sum	0	0	1	2	3

Subject 19					
	Response				
Day	0	1	2	3	Sum
1	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	1
2	0	0	0	0	0
3	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	1
4	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	1
Sum	1	1	0	1	3

- The contribution of subject 15 to Q_M is:

$$A_{15}(n_{15} - m_{15}) = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}'$$

$$A_{15}\Sigma_{15}A_{15}' = \begin{pmatrix} 0 & 0 & 0 \\ & 0.222 & -0.111 \\ & & 0.222 \end{pmatrix}$$

Accommodation of Missing Data

- The contribution of subject 19 to Q_M is:

$$A_{19}(n_{19} - m_{19}) = \begin{pmatrix} \frac{5}{3} & 0 & -\frac{1}{3} \end{pmatrix}'$$

$$A_{19}\Sigma_{19}A_{19}' = \begin{pmatrix} 1.556 & 0 & -0.778 \\ & 0 & 0 \\ & & 1.556 \end{pmatrix}$$

- With both complete and incomplete cases:

$$M = \begin{pmatrix} 4.5 \\ 0.5 \\ 0.5 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} + \begin{pmatrix} \frac{5}{3} \\ 0 \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 6.1667 \\ 0.8333 \\ 0.5000 \end{pmatrix}$$

$$V_M = \begin{pmatrix} 7.750 & -2.583 & -2.583 \\ & 7.750 & -2.583 \\ & & 7.750 \end{pmatrix} +$$

$$\begin{pmatrix} 0 & 0 & 0 \\ & 0.222 & -0.111 \\ & & 0.222 \end{pmatrix} + \begin{pmatrix} 1.556 & 0 & -0.778 \\ & 0 & 0 \\ & & 1.556 \end{pmatrix}$$

$$= \begin{pmatrix} 9.306 & -2.583 & -3.361 \\ & 7.972 & -2.694 \\ & & 9.528 \end{pmatrix}$$

Accommodation of Missing Data

- The mean score statistic with complete and incomplete cases is

$$Q_M = M' V_M^{-1} M = 7.441$$

with 3 df (p=0.059)

- The corresponding results using rank and ridit scores are:

	Complete Data		All Data	
Scores	Q_M	p	Q_M	p
Table	4.935	0.177	7.441	0.059
Rank	3.350	0.341	5.026	0.170
Ridit	3.350	0.341	5.497	0.139
Mod. Ridit	3.350	0.341	5.385	0.146

- Why are the results for the three types of rank scores not equal when all data are used?

Accommodation of Missing Data

- The two subjects with incomplete data can also be used in computing Q_C

- First, the correlation statistic for the complete cases is computed as follows (using the scores 1–4 for time and 0–3 for symptoms):

$$A_h = (\,1\;\;2\;\;3\;\;4\,) \otimes (\,0\;\;1\;\;2\;\;3\,) \text{ is } 1 \times 16$$

$$C_h = A_h(n_h - m_h) \text{ is a scalar}$$

$$C = \sum_{h=1}^{30} C_h = -15$$

$$V_C = \sum_{h=1}^{30} A_h \Sigma_h A_h' = 51.667$$

$$Q_C = C' V_C^{-1} C = (-15)^2/51.667 = 4.355$$

- With respect to the χ^2_1 distribution, p=0.037

Accommodation of Missing Data

- The contributions of subjects 15 and 19 are:

$$A_{15}(n_{15} - m_{15}) = -1 \qquad A_{15}\Sigma_{15}A'_{15} = 0.667$$

$$A_{19}(n_{19} - m_{19}) = -4.67 \qquad A_{19}\Sigma_{19}A'_{19} = 10.889$$

- With both complete and incomplete cases:

$$C = -15 - 1 - 4.667 = -20.667$$

$$V_C = 51.667 + 0.667 + 10.889 = 63.222$$

$$Q_C = (-20.667)^2/63.222 = 6.756 \qquad (p = 0.009)$$

- Corresponding results using rank & ridit scores:

	Complete Data		All Data	
Scores	Q_C	p	Q_C	p
Table	4.355	0.037	6.756	0.009
Rank	2.682	0.101	3.748	0.053
Ridit	2.682	0.101	4.494	0.034
Mod. Ridit	2.682	0.101	4.299	0.038

Relationship Between Q_C and Pearson's r

- Each subject's contribution to Q_C is related to Pearson's r between the row variable and the column variable

- The three (time, severity) pairs for subject 15 are (2,3), (3,3), and (4,2)

$$r = -1/\sqrt{4/3} = -0.866$$

- The (time, severity) pairs for subject 19 are (1,3), (3,1), and (4,0)

$$r = -4.667/\sqrt{21.778} = -1$$

- General results between Q_C and r:

$$A_h(n_h - m_h) = \text{numerator of } r$$

$$A_h \Sigma_h A_h' = \left(\frac{\text{denominator of } r}{\sqrt{n_h - 1}} \right)^2$$

Use of Q_M and Q_C for Continuous Data

- The randomization model tests were developed for stratified two-way contingency tables
- Q_M & Q_C can also be used for continuous data

Procedure:

- Let c denote the total number of observed values of the response
- Create a $t \times c$ contingency table for each subject
  - there will be one count of 1 and $c - 1$ counts of 0 in each of the t rows of the table
- Q_M tests if the mean scores across the t time points are equal
- Q_C tests if there is a linear association between time and response

Example

- In a dental study, the height of the ramus bone (mm) was measured in 20 boys at ages 8, $8\frac{1}{2}$, 9, and $9\frac{1}{2}$ years
- Two questions:
  - Does bone height change with age?
  - Is the association linear?
- If the assumptions of normal-theory methods are not justified, Q_M and Q_C can be used
- Since the response variable has 57 unique values, each subject has an underlying 4×57 contingency table

Reference

Elston, R. C. and Grizzle, J. E. (1962). Estimation of time-response curves and their confidence bands. *Biometrics* **18**, 148–159.

Data from Example

	Age (years)			
Subject	8	$8\frac{1}{2}$	9	$9\frac{1}{2}$
1	47.8	48.8	49.0	49.7
2	46.4	47.3	47.7	48.4
3	46.3	46.8	47.8	48.5
4	45.1	45.3	46.1	47.2
5	47.6	48.5	48.9	49.3
6	52.5	53.2	53.3	53.7
7	51.2	53.0	54.3	54.5
8	49.8	50.0	50.3	52.7
9	48.1	50.8	52.3	54.4
10	45.0	47.0	47.3	49.3
11	51.2	51.4	51.6	51.9
12	48.5	49.2	53.0	55.5
13	52.1	52.8	53.7	55.0
14	48.2	48.9	49.3	49.8
15	49.6	50.4	51.2	51.8
16	50.7	51.7	52.7	53.3
17	47.2	47.7	48.4	49.5
18	53.3	54.6	55.1	55.3
19	46.2	47.5	48.1	48.4
20	46.3	47.6	51.3	51.8

SAS Statements for Example

```
data a;
input subject h80 h85 h90 h95;
cards;
 1 47.8 48.8 49.0 49.7
       . . .
20 46.3 47.6 51.3 51.8
 ;
data b; set a;
keep subject age ramus;
age=8;   ramus=h80;  output;
age=8.5; ramus=h85;  output;
age=9;   ramus=h90;  output;
age=9.5; ramus=h95;  output;
proc freq;
tables subject*age*ramus / noprint cmh2;
```

- The `cmh2` option requests Q_M and Q_C only
- Q_M=41.293, df=3, $p < 0.001$
- Q_C=41.290, df=1, $p < 0.001$

Example

- Longitudinal study of 619 patients in 4 groups:
  1. kidney disease, hypertensive ($n = 294$)
  2. kidney disease, not hypertensive ($n = 103$)
  3. no kidney disease, hypertensive ($n = 73$)
  4. no kidney disease, not hypertensive ($n = 149$)

- Response variable is serum creatinine reciprocal (SCR), which ranges from 0.028 to 2.5

- Repeated measures factor is age (18–84 years)

- No. of observations/patient ranges from 1–22

- If normal-theory methods are not appropriate, Q_C can be used to test if there is a linear association between age and SCR in each group

Reference

Jones, R. H. and Boadi-Boateng, F. (1991). Unequally spaced longitudinal data with AR(1) serial correlation. *Biometrics* **47**, 161–175.

SAS Statements for Example

```
data a b c d; infile 'jbb.dat';
input id group age scr;
if group=1 then output a;
if group=2 then output b;
if group=3 then output c;
if group=4 then output d;
proc freq data=a;
tables id*age*scr / noprint cmh1;
proc freq data=b;
tables id*age*scr / noprint cmh1;
proc freq data=c;
tables id*age*scr / noprint cmh1;
proc freq data=d;
tables id*age*scr / noprint cmh1;
```

Results		
Group	Q_C	p-value
1	—[a]	—
2	2.80	0.094
3	4.68	0.031
4	7.31	0.007

[a]cannot be computed in group 1