January 13, 1997

# Analyzing Bivariate Continuous Data That Have Been Grouped Into Categories Defined by Sample Quantiles of the Marginal Distributions

Craig B. Borkowf[1], Mitchell H. Gail[1], Raymond Carroll[2] and Richard D. Gill[3]

[1] Biostatistics Branch, National Cancer Institute, Executive Plaza North, Room 431, 6130 Executive Blvd., MSC 7368, Rockville, Maryland 20892-7368, U.S.A.
TEL: (301) 496-3356, FAX: (301) 402-0081

[2] Department of Statistics, Texas A&M University, College Station, Texas 77843-3143

[3] Department of Mathematics, University of Utrecht, P.O.B. 80010, 3508 TA Utrecht, Netherlands

Key words: Agreement, bivariate sample quantile distribution, contingency table, kappa statistic, sample quantile

Summary

Epidemiologists sometimes study the association between two measures of exposure on the same subjects by grouping the data into categories that are defined by sample quantiles of the two marginal distributions. Although such grouped data are presented in a two-way contingency table, the cell counts in this table do not have a multinomial distribution. We use the term "bivariate quantile distribution" (BQD) to describe the joint distribution of counts in such a table. Blomqvist (1950) gave an exact BQD theory for the case of only 4 categories based on division at the sample medians. The asymptotic theory he presented was not valid, however, except in special cases. We present a valid asymptotic theory for arbitrary numbers of categories and apply this theory to construct confidence intervals for the kappa statistic. We show by simulations that the confidence interval procedures we propose have near nominal coverage for sample sizes exceeding 90, both for $2 \times 2$ and $3 \times 3$ tables. These simulations also illustrate that the asymptotic theory of Blomqvist (1950) and the methods given by Fleiss, Cohen and Everitt (1969) for multinomial sampling can yield subnominal coverage for BQD data, although in some cases the coverage for these procedures is near nominal levels.

# 1 Introduction

Epidemiologists sometimes cross-classify continuous bivariate data by determining the sample quantiles of each marginal distribution and categorizing the bivariate data into cells determined by these sample quantiles. For example, Pietinen, Hartman, Haapa et al. (1988a, 1988b) used sample quintiles (Table 1) to compare data on vitamin E consumption measured by two approaches: (1) two self-administered food frequency questionnaires based on recall of diet for the previous 12 months; and (2) a detailed prospective food consumption record covering 12 two-day periods spaced over a 6 month interval. The vitamin E consumption from the two food frequency questionnaires was computed as the mean of the two measurements which straddled the interval during which prospective food consumption records were taken.

To measure agreement between these two types of measurements, Pietinen, Hartman, Haapa, et al. (1988a, 1988b) used Pearson correlation coefficients based on the underlying continuous measurements, but they also examined quantities based on the sample quantiles, such as the proportion of subjects whose vitamin E consumption was rated in the lowest quintile by the food frequency questionnaire among those in the lowest quintile based on the food consumption record.

In this paper we develop asymptotic distribution theory for the counts in tables like Table 1, and we use this theory to derive confidence interval procedures for one measure of agreement in contingency tables, the kappa statistic (see Chapter 13 in Fleiss (1981) and Landis and Koch (1977a, 1977b)). This theory can be used to derive the asymptotic distribution of other measures of agreement, such as the proportion of measurements classified in the same or adjacent quantile categories on the two measurements (e.g. Willett, Sampson, Stampfer et al., 1985).

The counts in tables like Table 1 do not have a multinomial distribution because the cut-points used to classify the data are based on the sample quantiles, rather than fixed *a* *priori.* In particular, using sample quantiles as cut-points fixes the margins of the table, except for rounding, as illustrated in Table 1. We call the distribution of counts in such tables the "bivariate quantile distribution" (BQD).

Blomqvist (1950) derived the exact theory for the BQD for $2 \times 2$ tables partitioned at the sample medians. His asymptotic results are only valid under special assumptions, however.

We present notation and assumptions (Section 2) and derive asymptotic theory for the BQD (Section 3). We derive asymptotic theory and methods for construction of confidence intervals for kappa (Section 4). We present simulations to study the coverage of such confidence intervals (Section 5), not only for methods based on the BQD, but for the methods of Fleiss, Cohen and Everitt (1969), which are appropriate for multinomial sampling, and for a generalization of the results of Blomqvist (1950) to tables with more than four categories. We compare these methods on the data in Table 1 (Section 6) before discussing our results (Section 7).

## 2 Notation and Assumptions

Let the bivariate sample $(X_k, Y_k)$ for $k = 1, 2, ..., t$ be i.i.d. from the distribution $F$. Let $F$ have marginal distributions $G(x)$ and $H(y)$ and conditional distributions $G(x|y)$ and $H(y|x)$. Also, let $F(x, y)$ be differentiable as a function of $(x, y)$ at the quantiles of $G$ and $H$. That is, $\{F(x + h_1\Delta, y + h_2\Delta) - F(x, y)\}/\Delta = h_1 \frac{\partial}{\partial x} F(x, y) + h_2 \frac{\partial}{\partial y} F(x, y) + o(\Delta)$ uniformly in the direction vector $(h_1, h_2)$. The term $o(\Delta)$ is such that $o(\Delta)/\Delta$ tends to zero as $\Delta$ tends to zero. Let $\hat{F}(x, y)$, $\hat{G}(x)$, and $\hat{H}(y)$ denote the corresponding right-continuous empirical distribution (EDFs). For example, using the indicator function $I\{\cdot\}$, define

$$\hat{F}(x, y) = t^{-1} \sum_{k=1}^{t} I\{X_k \leq x, Y_k \leq y\}. \tag{2.1}$$

To categorize the $(X, Y)$ data into $r$ row and $c$ column categories, choose an increasing set of marginal proportions $\{\gamma_i\}$ $(i = 0, 1, ..., r)$ and $\{\eta_j\}$ $(j = 0, 1, ..., c)$ such that $\gamma_0 = \eta_0 = 0$ and $\gamma_r = \eta_c = 1$. We will concentrate on evenly spaced quantiles, $\gamma_i = i/r$ and $\eta_j = j/c$. For example, for quintiles, $r = c = 5$, $\gamma_i = i/5$, and $\eta_j = j/5$. The corresponding population $\gamma_i$-quantiles for $X$ are $\xi_i = G^{-1}(\gamma_i)$, and the population $\eta_j$-quantiles for $Y$ are $\psi_j = H^{-1}(\eta_j)$. We assume that $g(x) = G'(x)$ and $h(y) = H'(y)$ exist and are positive at the selected population quantiles, so these inverses are unique at the selected quantiles. For completeness, let $\xi_0 = \psi_0 = -\infty$ and $\xi_r = \psi_c = +\infty$. We also need the following notation based on these population quantiles: $\phi_{ij} = F(\xi_i, \psi_j)$, $\gamma_{i|j} = G(\xi_i|\psi_j)$, and $\eta_{j|i} = H(\psi_j|\xi_i)$. The parameters $\gamma_{i|j}$ and $\eta_{j|i}$ are crucial determinants of the asymptotic covariance structure. Sample estimates of these quantiles are given by the left-continuous quantities $u_i = \inf\{u : \gamma_i \leq \hat{G}(u)\}$ for $\xi_i$ and $v_j = \inf\{v : \eta_j \leq \hat{H}(v)\}$ for $\psi_j$ (Csörgő, 1983). For completeness, let $u_0 = v_0 = -\infty$ and $u_r = v_c = +\infty$.

The proportion of counts falling in the $(i, j)^{\text{th}}$ classification defined by $u_{i-1} < x \leq u_i$ and $v_{j-1} < y \leq v_j$ is

$$p_{ij} = \hat{F}(u_i, v_j) - \hat{F}(u_{i-1}, v_j) - \hat{F}(u_i, v_{j-1}) + \hat{F}(u_{i-1}, v_{j-1}). \tag{2.2}$$

Note, for example, that $p_{11}$ is the proportion of counts in the lowest quantiles of $X$ and $Y$ and corresponds to the "upper left" cell of the table, as in Table 1 and Figure 3. Thus, the cell counts in the $r \times c$ BQD table are given by $\{p_{ij}t\}$ $(i = 1, 2, ..., r;\ j = 1, 2, ..., c)$. As $t$ increases, the quantities $p_{ij}$ tend to

$$\pi_{ij} = F(\xi_i, \psi_j) - F(\xi_{i-1}, \psi_j) - F(\xi_i, \psi_{j-1}) + F(\xi_{i-1}, \psi_{j-1}). \tag{2.3}$$

We replace a subscript by a plus sign to indicate summation over that subscript. For example, $p_{i+} = \sum_{j=1}^{c} p_{ij}$. Note that $p_{i+} = \pi_{i+} = \gamma_i - \gamma_{i-1}$ and $p_{+j} = \pi_{+j} = \eta_j - \eta_{j-1}$.

# 3 The Bivariate Quantile Distribution

## 3.1 Asymptotic Theory

We use the fact (equation 2.2) that the sample proportions, $\{p_{ij}\}$, are linear combinations of the joint EDFs evaluated at the sample quantiles, $\left\{\hat{F}(u_i, v_j)\right\}$, and the expectations and covariances of the limiting distribution of $\left\{\hat{F}(u_i, v_j)\right\}$, to compute the asymptotic normal distribution for $\{p_{ij}\}$. In turn, we can use the delta-method to approximate the variance of measures of association that are functions of $\{p_{ij}\}$, such as the kappa statistic.

We approximate $\hat{F}(u_i, v_j)$ in terms of $\hat{F}(\xi_i, \psi_j)$, $\hat{G}(\xi_i)$, and $\hat{H}(\psi_j)$. Consider the decomposition

$$t^{\frac{1}{2}}\left\{\hat{F}(u_i, v_j) - F(\xi_i, \psi_j)\right\} = t^{\frac{1}{2}}\left\{\hat{F}(u_i, v_j) - F(u_i, v_j)\right\} + t^{\frac{1}{2}}\left\{F(u_i, v_j) - F(\xi_i, \psi_j)\right\}.$$

By the delta method, the second term on the right converges in distribution to $t^{\frac{1}{2}}\left\{\frac{\partial}{\partial x}F(\xi_i, \psi_j)(u_i - \xi_i) + \frac{\partial}{\partial y}F(\xi_i, \psi_j)(v_j - \psi_j)\right\}$ provided $F$ is differentiable at the quantiles. The first term on the right converges in distribution to $t^{\frac{1}{2}}\left\{\hat{F}(\xi_i, \psi_j) - F(\xi_i, \psi_j)\right\}$. To show this result, note that $u_i$ and $v_j$ converge in probability to $\xi_i$ and $\psi_j$, and the continuity of $F$ at $(\xi_i, \psi_j)$ ensures the continuity of limiting sample paths of $t^{\frac{1}{2}}\left\{\hat{F}(x, y) - F(x, y)\right\}$ at $(\xi_i, \psi_j)$. Adding $t^{\frac{1}{2}}F(\xi_i, \psi_j)$ to both sides of the decomposition and then dividing by $t^{\frac{1}{2}}$, we obtain the representation:

$$\hat{F}(u_i, v_j) = \hat{F}(\xi_i, \psi_j) + \frac{\partial}{\partial x}F(\xi_i, \psi_j)(u_i - \xi_i) + \frac{\partial}{\partial y}F(\xi_i, \psi_j)(v_j - \psi_j) + o_p\left(t^{-\frac{1}{2}}\right). \tag{3.1}$$

The notation $o_p\left(t^{-\frac{1}{2}}\right)$ means that the remainder term is stochastically negligible, namely $t^{\frac{1}{2}}o_p(t^{-\frac{1}{2}})$ converges to zero in probability (see Bishop, Feinberg, Holland, 1975, page 475).

Because the joint distribution function can be written as

$$F(x, y) = \int_{-\infty}^{x} H(y|z)g(z)dz = \int_{-\infty}^{y} G(x|z)h(z)dz, \tag{3.2}$$

we differentiate equation (3.2) with respect to $x$ and $y$ and evaluate at $(\xi_i, \psi_j)$ to obtain

$$\frac{\partial}{\partial x}F(\xi_i, \psi_j) = H(\psi_j|\xi_i)g(\xi_i) = \eta_{j|i}g(\xi_i) \tag{3.3}$$

and

$$\frac{\partial}{\partial y}F(\xi_i,\psi_j) = G(\xi_i|\psi_j)h(\psi_j) = \gamma_{i|j}h(\psi_j) \tag{3.4}$$

Because $G$ and $H$ are differentiable at the quantiles, the results of Ghosh (1971) and Gill (1989) yield the Bahadur representation,

$$(u_i - \xi_i) = [\gamma_i - \hat{G}(\xi_i)]/g(\xi_i) + o_p(t^{-\frac{1}{2}}) \tag{3.5}$$

and

$$(v_j - \psi_j) = [\eta_j - \hat{H}(\psi_j)]/h(\psi_j) + o_p(t^{-\frac{1}{2}}). \tag{3.6}$$

Substituting equations (3.3) through (3.6) into (3.1), we obtain

$$\hat{F}(u_i,v_j) = \hat{F}(\xi_i,\psi_j) - \eta_{j|i}\left[\hat{G}(\xi_i) - \gamma_i\right] - \gamma_{i|j}\left[\hat{H}(\psi_j) - \eta_j\right] + o_p(t^{-\frac{1}{2}}). \tag{3.7}$$

Because $t^{\frac{1}{2}}(\hat{F} - F, \hat{G} - G, \hat{H} - H)$ are jointly asymptotically normal, equation (3.7) implies that $t^{\frac{1}{2}}\left\{\hat{F}(u_i,v_j) - F(u_i,v_j)\right\}$ tends to normality, and, indeed, $t^{\frac{1}{2}}\left\{\hat{F}(u_i,v_j) - F(u_i,v_j)\right\}_{ij}$ tend to a jointly normal distribution.

To facilitate calculations, we define the vectors $\lambda'_{ij} = \begin{pmatrix} 1, & -\eta_{j|i}, & -\gamma_{i|j} \end{pmatrix}$, $\mu'_{ij} = \begin{pmatrix} 0, & \gamma_i, & \eta_j \end{pmatrix}$, and $w'_{ij} = \begin{Bmatrix} \hat{F}(\xi_i,\psi_j), & \hat{G}(\xi_i), & \hat{H}(\psi_j) \end{Bmatrix}$. Then, we can rewrite equation (3.7) as

$$\hat{F}(u_i,v_j) = \lambda'_{ij}(w_{ij} - \mu_{ij}) + o_p(t^{-\frac{1}{2}}). \tag{3.8}$$

Having approximated $\hat{F}(u_i,v_j)$ as a linear function of the EDFs evaluated at the population quantiles, we know it has a limiting normal distribution whose means and variances can be computed. Define $\theta'_{ij} = \begin{pmatrix} \phi_{ij} & \gamma_i & \eta_j \end{pmatrix}$, and let $m = \min\{i,k\}$ and $n = \min\{j,l\}$. Standard calculations (Appendix A) show that for every sample size $t$

$$E\left\{\hat{F}(\xi_i,\psi_j)\right\} = \phi_{ij} \tag{3.9}$$

and

$$\mathrm{Cov}\left\{t^{\frac{1}{2}}\hat{F}(\xi_i,\psi_j), t^{\frac{1}{2}}\hat{F}(\xi_k,\psi_l)\right\} = (\phi_{mn} - \phi_{ij}\phi_{kl}). \tag{3.10}$$

Note also that $G(x) = F(x,\infty)$ and $H(y) = F(\infty, y)$, which imply $\gamma_i = \phi_{ic}$ and $\eta_j = \phi_{rj}$. Thus from equation (3.10) we obtain

$$\mathrm{Cov}\left[t^{\frac{1}{2}}w_{ij}, t^{\frac{1}{2}}w_{kl}\right] = \begin{bmatrix} (\phi_{mn} - \phi_{ij}\phi_{kl}) & (\phi_{mj} - \phi_{ij}\gamma_k) & (\phi_{in} - \phi_{ij}\eta_l) \\ (\phi_{ml} - \gamma_i\phi_{kl}) & (\gamma_m - \gamma_i\gamma_k) & (\phi_{il} - \gamma_i\eta_l) \\ (\phi_{kn} - \eta_j\phi_{kl}) & (\phi_{kj} - \eta_j\gamma_k) & (\eta_n - \eta_j\eta_l) \end{bmatrix} \equiv \sum_{ijkl}. \tag{3.11}$$

It follows from equations (3.8), (3.9), and (3.11) that $t^{\frac{1}{2}}\left\{\hat{F}(u_i,v_j) - \phi_{ij}\right\}$ and $t^{\frac{1}{2}}\left\{\hat{F}(u_k,v_l) - \phi_{kl}\right\}$ are jointly asymptotically normal with mean zero and covariance

$$\lambda'_{ij}\sum_{ijkl}\lambda_{kl}. \tag{3.12}$$

In particular, the limiting variance of $t^{\frac{1}{2}}\left\{\hat{F}(u_i,v_j) - \phi_{ij}\right\}$ may be written without matrix notation as

$$\begin{aligned} & \phi_{ij}(1-\phi_{ij}) + \eta_{j|i}^2\gamma_i(1-\gamma_i) + \gamma_{i|j}^2\eta_j(1-\eta_j) \\ - \quad & 2\eta_{j|i}\phi_{ij}(1-\gamma_i) - 2\gamma_{i|j}\phi_{ij}(1-\eta_j) + 2\eta_{j|i}\gamma_{i|j}(\phi_{ij} - \gamma_i\eta_j). \end{aligned} \tag{3.13}$$

For most applications, the variances and covariances involve so many terms that matrix notation and computer calculations are needed.

## 3.2 Parameter Estimation

To estimate the covariances of $\{p_{ij}\}$ from equations (2.2), (3.8) and (3.11), we need to estimate $\{\phi_{ij}\}$, $\{\gamma_{i|j}\}$, and $\{\eta_{j|i}\}$. We estimate $\{\phi_{ij}\}$ by $\left\{\hat{F}(u_i,v_j)\right\}$.

The estimation of $\{\gamma_{i|j}\}$ and $\{\eta_{j|i}\}$ is difficult because, for example, $\gamma_{i|j} = P(X \le \xi_i | Y = \psi_j)$, and in finite samples there will be no pairs $(X,Y)$ with $Y = \psi_j$ exactly. Thus some kind of smoothing procedure is needed, analogous to density estimation (Silverman, 1986). Our estimate is based on

$$\gamma_{i|j} = G(\xi_i|\psi_j) = P\left\{X \le \xi_i | Y = \psi_j\right\} = P\left\{G(X) \le \gamma_i | H(Y) = \eta_j\right\}, \tag{3.14}$$

which leads to

$$\begin{aligned}\hat{\gamma}_{i|j} &= \hat{P}\left\{\hat{G}(X_k) \le \gamma_i \,\Big|\; |\hat{H}(Y_k) - \left(\eta_j + \frac{1}{2t}\right)| \le \beta_t/t\right\} \\ &= \frac{\sum_{k=1}^{t} I\left\{\hat{G}(X_k) \le \gamma_i, |\hat{H}(Y_k) - (\eta_j + \frac{1}{2t})| \le \beta_t/t \right\}}{\sum_{k=1}^{t} I\left\{|\hat{H}(Y_k) - (\eta_j + \frac{1}{2t})| \le \beta_t/t\right\}}. \qquad (3.15)\end{aligned}$$

In an analogous fashion, we can define estimates $\left\{\hat{\eta}_{j|i}\right\}$ of $\left\{\eta_{j|i}\right\}$. To obtain consistent estimates of $\gamma_{i|j}$ and $\eta_{j|i}$, we require $\beta_t \to \infty$ and $\beta_t/t \to 0$ as $t \to \infty$. In $2 \times 2$ tables partitioned by medians, we use $\beta_t = (t/2)^{\frac{1}{2}}$ and in $3 \times 3$ tables partitioned by tertiles, we use $\beta_t = (t/3)^{\frac{1}{2}}$.

We could also estimate the covariances of $\{p_{ij}\}$ by a bootstrap procedure (Efron and Tibshirani, 1986). This procedure is valid under the same conditions needed for the asymptotic theory in Section 3.1, as follows from general results in Gill (1989).

# 4 The Kappa Statistic ($\kappa$), its Asymptotic Variance, and Confidence Intervals

## 4.1 The Kappa Statistic ($\kappa$)

The kappa statistic ($\kappa$) (Cohen, 1960) measures the agreement between two variables in $r \times r$ tables. This statistic was originally used in psychological studies with nominal categories and, thus, for counts following the multinomial distribution. Landis and Koch (1977a, 1977b) discuss the use of $\kappa$ for ordinal data and provide some useful benchmarks for its interpretation.

Let $\Pi_0 = \sum_{i=1}^{r} \pi_{ii}$ and $\Pi_e = \sum_{i=1}^{r} \pi_{i+}\pi_{+i}$. Then $\Pi_0$ represents the limiting proportion of diagonal observations, while $\Pi_e$ represents the limiting proportion of diagonal counts that we would expect if the underlying variates $X$ and $Y$ were independent. The quantity $\kappa$ is defined by

$$\kappa = \frac{\Pi_0 - \Pi_e}{1 - \Pi_e}. \qquad (4.1)$$

Note that $\kappa = 1$ corresponds to perfect agreement. The sample estimate of $\kappa$ is

$$\hat{\kappa} = \frac{P_0 - P_e}{1 - P_e}, \tag{4.2}$$

where $P_0$ and $P_e$ estimate $\Pi_0$ and $\Pi_e$ respectively, by replacing $\pi_{ij}$ by $p_{ij}$ in the defining formulas above.

Under bivariate quantile sampling, the marginal distributions are fixed, and $P_e = \Pi_e$. Therefore, we estimate the variance of $\kappa$ as

$$\hat{\text{V}}\text{ar}_{BQ}\,(\hat{\kappa}) = (1 - \Pi_e)^{-2} \sum_{i=1}^{r} \sum_{j=1}^{r} \hat{\text{C}}\text{ov}_{BQ}(p_{ii}, p_{jj}). \tag{4.3}$$

The needed $\hat{\text{C}}\text{ov}_{BQ}(p_{ii}, p_{jj})$ is obtained as in Section 3 under bivariate quantile sampling.

## 4.2 Estimates of the Variance of $\hat{\kappa}$ Under Other Sampling Models

For completeness, we compare variance estimates under BQD sampling with other estimates of the variance of $\hat{\kappa}$ appropriate for other sampling plans. Under multinomial sampling (MULT), the cell counts have random marginal totals, and both $P_0$ and $P_e$ are random variables. Fleiss, Cohen, and Everitt (1969) (FCE) used the delta-method to derive the estimated variance of $\hat{\kappa}$ for multinomial samples. Agresti (1990) presented an algebraically equivalent but computationally simpler asymptotic approximation, namely:

$$\hat{\text{V}}\text{ar}\,(\hat{\kappa}) = t^{-1} \left( \frac{P_0(1 - P_0)}{(1 - P_e)^2} + \frac{2(1 - P_0)C_1}{(1 - P_e)^3} + \frac{(1 - P_0)C_2}{(1 - P_e)^4} \right), \tag{4.4}$$

where $C_1 = 2P_0P_e - \sum_{i=1}^{r} p_{ii}(p_{i+} + p_{+i})$ and $C_2 = \sum_{i=1}^{r} \sum_{j=1}^{r} p_{ij}(p_{j+} + p_{+i})^2 - 4P_e^2$. We can replace $\{p_{ij}\}$ by $\{\pi_{ij}\}$ to obtain the asymptotic variance of $\hat{\kappa}$ for multinomial tables.

Blomqvist (1950) gave the following asymptotic formula for the variance of the $\{p_{ij}\}$ in $2 \times 2$ tables partitioned by sample medians:

$$\text{Var}\left(t^{\frac{1}{2}} p_{ij}\right) \to \pi_{11}\,(0.5 - \pi_{11})\,. \tag{4.5}$$

The following argument shows that this result corresponds to the asymptotic variance from the singular multivariate normal distribution to which the multivariate extended hypergeometric (MXH) distribution converges asymptotically. The multivariate extended hypergeometric distribution is obtained from an arbitrary multinomial distribution of counts in an $r \times c$ table by conditioning on the margins (Plackett, 1981, page 64). The term "extended" refers to the fact that cell means may differ from their expectations under independence. Under MXH sampling, the marginal counts are fixed, and $P_e = \Pi_e$. Therefore $\widehat{\text{Var}}_{MXH}(\hat{\kappa})$ can be estimated from equation (4.3) with multivariate extended hypergeometric covariances $\hat{\text{Cov}}_{MXH}(p_{ii}, p_{jj})$ in place of $\hat{\text{Cov}}_{BQ}(p_{ii}, p_{jj})$. The terms $\hat{\text{Cov}}_{MXH}(p_{ii}, p_{jj})$ may be estimated by substituting $p_{ij}$ for $\pi_{ij}$ in asymptotic expressions given by Plackett (1981, page 65). Plackett gives the asymptotic quadratic form in the normal approximation to the distribution of $\{p_{ij}\}$, from which required covariances can be calculated. In the special case of $2 \times 2$ tables under MXH sampling, $\text{Var}(p_{ij}) \to t^{-1}(\pi_{11}^{-1} + \pi_{12}^{-1} + \pi_{21}^{-1} + \pi_{22}^{-1})^{-1}$, which reduces to equation (4.5) because $\pi_{11} = \pi_{22} = \frac{1}{2} - \pi_{12} = \frac{1}{2} - \pi_{21}$. To improve performance in sparse tables (Cox 1970, page 33), we substituted $p_{ij} + (4t)^{-1}$ for $\pi_{ij}$ in the formulas of Plackett, and we divided the resulting estimated asymptotic covariances of $t^{\frac{1}{2}} p_{ij}$ by $t - 1$, instead of by $t$, to obtain $\hat{\text{Cov}}_{MXH}(p_{ii}, p_{jj})$. Division by $(t-1)$ agrees with the exact calculation of $\text{Cov}(p_{ii}, p_{jj})$ under MXH sampling when $X$ and $Y$ are independent.

The limiting normal distribution theory for BQD sampling and MXH sampling agree under certain conditions, defined in Corollaries 1 and 2 and Theorem 2.

**Theorem 1**: The quantities $t^{\frac{1}{2}} \hat{F}(u_i, v_j)$ and $t^{\frac{1}{2}} \hat{F}(u_k, v_\ell)$ have the same limiting covariance under BQD and MXH sampling if $\gamma_{i|j} = \gamma_i$, $\eta_{j|i} = \eta_j$, $\phi_{ij} = \gamma_i \eta_j$, $\gamma_{k|\ell} = \gamma_k$, $\eta_{\ell|k} = \eta_\ell$ and $\phi_{k\ell} = \gamma_k \eta_\ell$. **Proof** is in Appendix B.

**Corollary 1**: $\left\{t^{\frac{1}{2}} p_{ij}\right\}$ have the same limiting variances and covariances under BQD and MXH sampling for all $i = 0, 1, ..., r$, $j = 0, 1, ..., c$ if $\gamma_{i|j} = \gamma_i$, $\eta_{j|i} = \eta_j$ and $\phi_{ij} = \gamma_i \eta_j$ for all $i = 0, 1, ..., r$ and $j = 0, 1, ..., c$. **Proof**: This Corollary follows from Theorem 1 and equation (2.2).

**Corollary 2**: If $X$ and $Y$ are independent, $\left\{t^{\frac{1}{2}} p_{ij}\right\}$ have the same limiting variances and covariances under BQD and MXH sampling. **Proof**: Independence implies $\gamma_{i|j} = \gamma_i$, $\eta_{j|i} = \eta_j$ and $\phi_{ij} = \gamma_i \eta_j$ for all $i$ and $j$. **Comment**: Independence of $X$ and $Y$ is a stronger condition than the conditions in Corollary 1, which only require that counts based on the cross-classification of $X$ and $Y$ according to the population quantiles be independent in the table defined by this cross-classification.

The conditions of Corollary 1 also apply to $2 \times 2$ tables. In the case $\eta_1 = \gamma_1 = 0.5$, corresponding to division at the medians, however, we have the following special result.

**Theorem 2**: For a $2 \times 2$ table with $\gamma_1 = \eta_1 = 0.5$, $\left\{t^{\frac{1}{2}} p_{ij}\right\}$ have the same limiting variances and covariances under BQD and MXH sampling if $\gamma_{1|1} = \eta_{1|1} = 0.5$. **Proof**: Under MXH sampling, the limiting variance of $t^{\frac{1}{2}} p_{ij}$ is $\pi_{11}(0.5 - \pi_{11}) = \phi_{11}(0.5 - \phi_{11})$ from equation (4.5). Under BQD sampling, substitution of $\gamma_{1|1} = \eta_{1|1} = \gamma_1 = \eta_1 = 0.5$ into equation (3.13) yields the same limiting variance. Because $p_{22} = p_{11}$, $p_{12} = 0.5 - p_{11}$ and $p_{21} = 0.5 - p_{11}$, all other limiting variances and covariances of $\left\{t^{\frac{1}{2}} p_{ij}\right\}$ are also equal under MXH and BQD sampling. **Comment**: Independence is not required for the conditions of Theorem 2 to hold. For example, the conditions hold for the bivariate normal distribution with non-zero correlation.

## 4.3 Confidence Interval Construction

We study confidence intervals $\hat{\kappa} \pm Z_{1-\alpha/2} \left\{\hat{\mathrm{V}}\mathrm{ar}(\hat{\kappa})\right\}^{-\frac{1}{2}}$, where $\hat{\mathrm{V}}\mathrm{ar}(\hat{\kappa})$ is estimated either under bivariate quantile, multinomial or multivariate extended hypergeometric sampling models, as in Section 3, and where $Z_{1-\alpha/2} = \Phi^{-1}(1 - \alpha/2)$ is the $1 - \alpha/2$ quantile of the standard normal distribution $\Phi$.

We also study confidence intervals based on the bootstrap algorithm (Efron and

Tibshirani, 1986). The validity of these procedures follows from general results in Gill (1989) under the same assumptions required for the asymptotic theory of Section 3.1. If $\hat{\kappa}^b$ represents an estimate of $\hat{\kappa}$ based on bootstrap replicate $b$, and if there are $B$ bootstrap replicates, then we compute a confidence interval from $\hat{\kappa} \pm 1.96s$, where $s^2 = \Sigma(\hat{\kappa}^b - \bar{\kappa})/(B-1)$ and $\bar{\kappa} = \Sigma\hat{\kappa}^b/B$. We describe this as the BSV procedure to indicate that it is based on the bootstrap sample variance. We also calculate a confidence interval $(\kappa_L, \kappa_U)$ where $\kappa_L$ and $\kappa_U$ are the $2.5^{\text{th}}$ and $97.5^{\text{th}}$ percentiles of the bootstrap sample. This confidence interval procedure is denoted BPC. Bootstrap samples are obtained by resampling $t$ pairs with replacement from the original sample $(X_k, Y_k)$, $k = 1, 2, ..., t$.

# 5 Simulations and Other Numerical Studies

## 5.1 Asymptotic Theory for Several Bivariate Distributions, $F(x, y)$

We consider several bivariate distributions to illustrate differences in asymptotic theory that arise under BQD sampling, multinomial sampling (MULT) and multivariate extended hypergeometric sampling (MXH). We let $\gamma_i = i/r$ and $\eta_j = j/r$ correspond to equal marginal proportions in an $r \times r$ table.

Bivariate normal distribution, BVN($\rho$).

The distribution is bivariate normal with means zero, variances 1 and correlation $\rho$. Unreported numerical studies by us confirm (see Theorem 2) that the asymptotic covariance of $t^{\frac{1}{2}}p_{ij}$ under BQD sampling equals that under MXH sampling for $2 \times 2$ tables based on medians, regardless of $\rho$. Note that $\gamma_1 = \gamma_{1|1} = \eta_1 = \eta_{1|1} = 0.5$, regardless of $\rho$ in the $2 \times 2$ case (Table 2, part a). Likewise, unreported numerical studies confirm the result of Corollary 2 for $3 \times 3$ tables with $\gamma_1 = \gamma_{1|1} = \gamma_{1|2} = \frac{1}{3}$, $\gamma_2 = \gamma_{2|1} = \gamma_{2|2} = \frac{2}{3}$ and $\rho = 0$ (Table 3, part a). The asymptotic covariances of $t^{\frac{1}{2}}p_{ij}$ under multinomial sampling differ from those under BQD and MXH sampling in all BVN($\rho$) cases and in all other cases described below.

Despite the fact that the counts have different asymptotic covariances under multinomial sampling, it is a mathematical coincidence that the limiting variance of $t^{\frac{1}{2}}\hat{\kappa}$ is the same for MULT, MXH and BQD sampling in $2 \times 2$ tables when the underlying distribution is BVN($\rho$) (Table 2, part b). For $3 \times 3$ tables with $\rho \neq 0$, there are slight differences in the limiting variance of $t^{\frac{1}{2}}\hat{\kappa}$ for BQD, MULT and MXH sampling (Table 3, part b).

Bivariate chi-squared distribution, BCH($\rho$)

BCH($\rho$) data are obtained by generating pairs from BVN($\rho^{\frac{1}{2}}$) and squaring each component. The marginal distributions $G(x)$ and $H(y)$ are each chi-square, and independence corresponds to $\rho = 0$.

For $\rho = 0$, Corollary 2 applies, and the asymptotic covariances of $t^{\frac{1}{2}}p_{ij}$ are equal for BQD and MXH sampling, but not for MULT sampling, both for $2 \times 2$ and $3 \times 3$ tables. For $\rho = 0.5$ or $0.9$, $\gamma_{i|j} \neq \gamma_i$ (Tables 2 and 3, part a), and the asymptotic covariances of $t^{\frac{1}{2}}p_{ij}$ differ for BQD, MXH and MULT sampling.

For $\rho = 0$, the asymptotic variances of $t^{\frac{1}{2}}\hat{\kappa}$ are equal for all three sampling schemes (data not shown), but slight differences are present with $\rho = 0.5$ or $0.9$ for $2 \times 2$ (Table 2, part b), and $3 \times 3$ (Table 3, part b) tables.

Nicked square distribution, NS

The NS has density 1 in the grey region of Figure 1, 2 in the black region of Figure 1 and 0 in the white region. Note that $Y$ and $X$ are dependent but uncorrelated.

For $2 \times 2$ tables, $\gamma_{1|1} = P(X \leq \frac{1}{2}|Y = \frac{1}{2}) = 0.5$, but $\eta_{1|1} = P(Y \leq \frac{1}{2}|X = \frac{1}{2}) = 0$ (Table 2, part a). The asymptotic covariances of $t^{\frac{1}{2}}p_{ij}$ differ for BQD, MXH and MULT sampling in this case, and the limiting variance of $t^{\frac{1}{2}}\hat{\kappa}$ is 2.0 for BQD sampling and 1.0 for MXH and MULT sampling (Table 2, part b). For $3 \times 3$ tables, $\gamma_{i|j} = \gamma_i$, $\eta_{j|i} = \eta_j$ and $\phi_{ij} = \gamma_i\eta_j$ (Table 3, part a). In this case, the asymptotic covariances of $t^{\frac{1}{2}}p_{ij}$ agree for BQD and MXH sampling (Corollary 1), but not for MULT sampling. Nonetheless, the asymptotic variances of $t^{\frac{1}{2}}\hat{\kappa}$ are the same under all three sampling plans (Table 3, part b).

Three squares distribution, TS

The TS distribution has density equal to 3 in the dark squares (Figure 2) and zero elsewhere. The lower left dark square is $\left[0, \frac{1}{3}\right] \times \left[0, \frac{1}{3}\right]$, the middle right dark square is $\left(\frac{2}{3}, 1\right] \times \left(\frac{1}{3}, \frac{2}{3}\right]$, and the remaining dark square is $\left(\frac{1}{3}, \frac{2}{3}\right] \times \left(\frac{2}{3}, 1\right]$. The variates $Y$ and $X$ are each uniformly distributed on $[0, 1]$, but $Y$ and $X$ are dependent, with covariance .280704-.25=.03704 and correlation .4444.

For $2 \times 2$ tables, $\gamma_{1|1} = 0$ and $\eta_{1|1} = 0$ (Table 2, part a), and the limiting covariances of $t^{\frac{1}{2}}p_{ij}$ differ for BQD, MXH and MULT sampling. The limiting variance of $t^{\frac{1}{2}}\hat{\kappa}$ is four times as great under BQD sampling as under MXH and MULT sampling (Table 2, part b).

For $3 \times 3$ tables, $\gamma_{1|2} = \eta_{1|2} = 0$ (Table 3, part a). However, $\xi_1 = \psi_1 = \frac{1}{3}$, and $G(\xi_1|y)$ is discontinuous in $y$ at $y = \frac{1}{3}$. Likewise, $H(\psi_1|x)$ is discontinuous in $x$ at $x = \frac{1}{3}$. Similarly for $\xi_2 = \psi_2 = \frac{2}{3}$, $G(\xi_2|y)$ is discontinuous in $y$ at $y = \frac{1}{3}$ and $y = \frac{2}{3}$ and $H(\psi_2|x)$ is discontinuous in $x$ at $x = \frac{1}{3}$ and $\frac{2}{3}$. Thus, the conditional probabilities $\gamma_{1|1}$, $\gamma_{2|1}$, $\gamma_{2|2}$, $\eta_{1|1}$, $\eta_{2|1}$ and $\eta_{2|2}$ are not defined (Table 3, part a). It follows that expressions (3.3) and (3.4) are not defined, and the variances and covariances of $t^{\frac{1}{2}}p_{ij}$ under BQD sampling cannot be determined (Table 3, part b ) by the methods of Section 3.1. Under MXH sampling, the limiting variances and covariances of $t^{\frac{1}{2}}p_{ij}$ are all zero. The limiting variances of $t^{\frac{1}{2}}\hat{\kappa}$ are 0.5 and 0.0 under MULT and MXH sampling, respectively, and undetermined for BQD sampling (Table 3, part b).

## 5.2 Simulation Studies of the Variances of $t^{\frac{1}{2}}\hat{\kappa}$ From Finite BQD Samples

We undertook simulation studies under BQD sampling to determine how large sample sizes must be for asymptotic BQD variance calculations to yield reliable results for $t^{\frac{1}{2}}\hat{\kappa}$ and to verify that asymptotic variance calculations under MXH and MULT sampling can be incorrect.

Random numbers were generated in the GAUSS 3.0 programming language (Aptech Systems, Inc., 1992) using the procedure RNDNS, an acceptance-rejection algorithm, for BVN($\rho$) variates and the procedure RNDUS, a multiplicative-congruential algorithm, for uniform variates. Normal variates were used to generate BVN($\rho$) and BCH($\rho$) data, as described in Section 5.1, and uniform variates were used to generate NS and TS data. The estimated variance of the quantity $t^{\frac{1}{2}}\hat{\kappa}_i = a_i$, from simulation $i$, based on $n$=100,000 simulations, was $s^2 = (n-1)^{-1}\Sigma(a_i - \bar{a})^2$, where $\bar{a} = \Sigma a_i/n$ and summations range from $i = 1$ to $i = n \equiv 100,000$. Each column in Table 2, part c and Table 3, part c required about 5 hours of computing time on a 90 MHz Pentium$^{\mathrm{TM}}$ processor. Each entry in part c of Tables 1 and 2 is independent of other entries.

For $2 \times 2$ tables, the sample variance is within 5% of the BQD asymptotic variance for sample size $t$=30 for all BVN($\rho$) distributions and for BCH(.5) (Table 2, part c). For BCH(.9) data, a sample size of $t$=60 is sufficient to bring the sample variance within 5% of the asymptotic variance. That is, (.5931-.5707)$\times$100/.5707=3.9%. Likewise, for $t$=60, the sample variance is only 1.7% smaller than the asymptotic variance for TS data. For NS data, the sample variance remains 10.3% smaller than the asymptotic variance, 2.0, even for $t$=1200, and for smaller sample sizes the asymptotic variance seriously overestimates the actual variance under BQD sampling.

The asymptotic variances computed under MULT and MXH sampling differ significantly from sample variances with $t$=1200 for BCH(.5), BCH(.9), NS and TS data. These are cases in which the asymptotic variances under MULT and MXH sampling differ from the BQD asymptotic variance. Assuming $s^2(n-1) \sim \sigma^2\chi^2_{n-1}$, where $\sigma^2 = \mathrm{Var}(t^{\frac{1}{2}}\hat{\kappa})$ and that $n$ is large enough so that the chi-square distribution is well approximated by normality, we can test whether the quantity $\sigma^2$ equals the asymptotic variance computed in Table 2, part b, using the standard normal deviate $Z = (s^2/\sigma^2 - 1)\{(n-1)/2\}^{\frac{1}{2}}$. For example, for BCH(.5) data and $\sigma^2 = .9355$ computed under MXH sampling, $z = (.9681/.9355 - 1)\,(99999/2)^{\frac{1}{2}} = 7.79$, giving strong evidence that the MXH calculation (and the identical MULT calculation) are misleading. These deviations are even more obvious for BCH(.9), NS and TS data.

Very similar results are obtained for $3 \times 3$ tables, except that the BQD asymptotic variance is close to the sample variance for NS data, even with $t = 30$ (Table 3, part c). Moreover, sample variances from BVN($\rho$) data differ significantly from asymptotic variances computed for MXH sampling when $\rho \neq 0$.

## 5.3 Simulated Coverage Under BQD Sampling

We simulated data under BQD sampling to assess the coverage of various procedures for constructing nominal 95% confidence intervals on $\kappa$. The same simulated data were analyzed by each procedure to facilitate comparisons. Results are based on 100,000 simulated trials except for the bootstrap procedures BSV and BPC, for which 1,000 trials and $B$=400 bootstrap repetitions were used.

For $2 \times 2$ tables (Table 4), the BQD procedure (see Sections 3.2 and 4.3) has near nominal size for sample sizes $t$ of 90 or more, except for the TS distribution, for which a sample size of 150 yields near nominal coverage. The BSV procedure performs similarly to the BQD procedure, although the BSV coverage is appreciably higher than 0.95 for small sample sizes. The BPC procedure has coverage consistently above nominal levels, even for $t$=300, except for the TS distribution. Simulations with 1,000 trials and with $t$=3,000 from the BVN(0) distribution yield a coverage of .960 for the BPC procedure, 0.953 for BSV and 0.952 for BQD.

The MXH and MULT procedures are identical for $2 \times 2$ tables (see Table 2, part b). For samples of $t$=150 or more and for all BVN($\rho$) distributions, for which these procedures have the appropriate asymptotic variance, coverage is near nominal levels (Table 4). Even for distributions such as BCH(.9) for which these procedures have inappropriate variances, the coverage is near nominal levels for $t \geq 150$. The coverage is substantially less than nominal, however, for the NS and TS distributions, for which the MXH and MULT sampling theory yields misleading results under BQD sampling.

Similar results were found for $3 \times 3$ tables (Table 5), except that the BQD procedures performs well even for the TS distribution, for which the variance is ill-defined, and the BSV bootstrap procedure tends to have supranominal coverage even for $t$=300. Both the MXH and MULT procedures have near nominal coverage for $t \geq 150$ for all distributions except the TS distribution.

To summarize, the BQD procedure yields near nominal coverage under BQD sampling for sample sizes above 90, and the bootstrap procedure BSV also works well for slightly larger sample sizes. The BPC procedure tends to have supranominal coverage in these simulations. The MXH and MULT procedures perform well except for distributions such as NS and TS, for which the MXH and MULT asymptotic theory is quite misleading under BQD sampling.

# 6 Example

We estimated $\hat{\kappa}$=0.2147 from Table 1 and obtained estimated standard deviations of $\hat{\kappa}$ of 0.0515, 0.0483 and 0.0473 respectively from BQD, MULT and MXH procedures. The estimated standard deviation of $\hat{\kappa}$ is 0.0500 if a different bandwidth, $2 \times (t/5)^{\frac{1}{2}}$, is used instead of $(t/5)^{\frac{1}{2}}$. Confidence intervals for $\hat{\kappa}$ computed under the BQD, BSV, MXH and MULT procedures were, respectively, (.1137,.3157), (.1068,.3226), (.1220,.3074) and (.1201,.3094).

Based on asymptotic theory and on the simulations in Section 5, we recommend the BQD procedure and confidence interval (.1137,.3157). It is reassuring, however, that discrepancies among these procedures are small.

# 7 Discussion

In this paper we develop the asymptotic theory for counts in a contingency table defined by BQD sampling. This theory extends and corrects the asymptotic theory given by Blomqvist (1950) for $2 \times 2$ tables, which is only correct for certain distributions, such as the bivariate normal distribution, that satisfy the conditions of Theorem 2.

This BQD asymptotic theory can be used to study many measures of association or agreement in BQD tables. We have focussed on the kappa statistic because of its frequent use, despite well known objections (Maclure and Willett, 1987). It is a mathematical coincidence that the asymptotic distribution of the estimate $\hat{\kappa}$ is the same under BQD, MXH and MULT sampling for $2 \times 2$ tables partitioned at the sample medians when the underlying data are bivariate normal. This result suggests, and our simulations confirm, that available confidence interval procedures for kappa constructed under multinomial sampling (Fleiss, Cohen and Everitt, 1969 and Agresti, 1990) will not be very misleading in many cases. We have constructed examples from non-normal distributions, however, for which the coverage of confidence intervals based on MXH or MULT sampling is below nominal levels. Therefore, we recommend the procedures developed for BQD sampling, or the bootstrap procedure, BSV, when the data arise by BQD sampling.

Further work might be useful to develop and evaluate alternative non-parametric estimators of parameters such as $\gamma_{i|j}$. We are currently developing parametric theory for BQD sampling to investigate issues of efficiency. Nonetheless, it is an attractive feature of the procedures presented in this paper that parametric assumptions are avoided.

**Acknowledgements**

We would like to thank Dr. Anne Hartman for bringing the problem of BQ tables in epidemiological studies to our attention. We are grateful to her and her collaborators for providing us with the data that appears in Table 1. We would also like to thank Drs. Tom Fears and Phil Rosenberg for statistical and computing assistance, and Mrs. Jennifer

Donaldson for typing the manuscript. Finally, Craig Borkowf, who is a graduate student at Cornell University, received support from the National Cancer Institute's Student Research Training Program grant #SRTP-E-600. Raymond Carroll was supported by a National Cancer Institute Grant (CA-57030). Carroll's research was partially completed while visiting the Institut für Statistik und Ökonometrie, Sonderforschungsbereich 373, Humboldt Universität zu Berlin, with partial support from a senior Alexander von Humboldt Foundation research award.

Appendix A. Covariance of $\hat{F}(\xi_i, \psi_j)$ and $\hat{F}(\xi_j, \psi_\ell)$

The expectation of $\hat{F}(\xi_i, \psi_j)$ is $E\left\{I(X \leq \xi_i, Y \leq \psi_j)\right\} = \phi_{ij}$. Likewise,

$$\begin{aligned}
\mathrm{Cov}\left\{\hat{F}(\xi_i, \psi_j), \hat{F}(\xi_k, \psi_\ell)\right\} &= t^{-2}\mathrm{Cov}\left\{\sum_{a=1}^{t} I(X_a \leq \xi_i, Y_a \leq \psi_j), \sum_{b=1}^{t} I(X_b \leq \xi_k, Y_b \leq \psi_\ell)\right\} \\
&= t^{-2}\sum_{a=1}^{t} \mathrm{Cov}\left\{I(X_a \leq \xi_i, Y_a \leq \psi_j), I(X_a \leq \xi_k, Y_a \leq \psi_\ell)\right\} \\
&= t^{-1}\left[E\left\{I(X_a \leq \xi_i, Y_a \leq \psi_j) I(X_a \leq \xi_k, Y_a \leq \psi_\ell)\right\} - \phi_{ij}\phi_{k\ell}\right] \\
&= t^{-1}\left[\mathrm{Prob}\left\{X_a \leq \min(\xi_i, \xi_k), Y_a \leq \min(\psi_j, \psi_\ell)\right\} - \phi_{ij}\phi_{k\ell}\right].
\end{aligned}$$

Appendix B. Proof of Theorem 1

Under the assumptions of Theorem 1, expression (3.13) for the limiting covariance of $t^{\frac{1}{2}}\hat{F}(u_i, v_j)$ and $t^{\frac{1}{2}}\hat{F}(u_k, v_\ell)$ under BQD sampling reduces, after some algebra, to

$$(\gamma_m - \gamma_i\gamma_k)(\eta_n - \eta_j\eta_\ell) \tag{A2.1}$$

where $m = \min(i, k)$ and $n = \min(j, \ell)$.

Now consider a $3 \times 3$ table with fixed marginal counts as shown in Figure 3. From standard results (page 65 in Plackett, 1981) for the multivariate hypergeometric distribution under the independence condition $\phi_{ij} = \gamma_i\eta_j$, the limiting covariances of $t^{\frac{1}{2}}$ times the quantities $a$, $b$, $c$ and $d$ are:

$$\begin{aligned}
\mathrm{Cov}\left(t^{\frac{1}{2}}a, t^{-\frac{1}{2}}a\right) &\to \eta_1(1-\eta_1)\gamma_1(1-\gamma_1), \\
\mathrm{Cov}\left(t^{-\frac{1}{2}}a, t^{-\frac{1}{2}}b\right) &\to -\gamma_1(\gamma_2-\gamma_1)\eta_1(1-\eta_1) \\
\mathrm{Cov}\left(t^{-\frac{1}{2}}a, t^{-\frac{1}{2}}c\right) &\to -\gamma_1(1-\gamma_1)\eta_1(\eta_2-\eta_1) \text{ and} \\
\mathrm{Cov}\left(t^{-\frac{1}{2}}a, t^{-\frac{1}{2}}d\right) &= \mathrm{Cov}\left(t^{-\frac{1}{2}}b, t^{\frac{1}{2}}c\right) \to \gamma_1(\gamma_2-\gamma_1)\eta_1(\eta_2-\eta_1).
\end{aligned}$$