

Article

Families of Generalized Quasisymmetry Models: A ϕ -Divergence Approach

Maria Kateri 

Institute of Statistics, RWTH Aachen University, 52062 Aachen, Germany; maria.kateri@rwth-aachen.de

Abstract: The quasisymmetry (QS) model for square contingency tables is revisited, highlighting properties and features on the basis of its alternative definitions. More parsimonious QS-type models, such as the ordinal QS model for ordinal classification variables and models based on association models (AMs) with homogeneous row and column scores, are discussed. All these models are linked to the local odds ratios (LOR). QS-type models and AMs were extended in the literature for generalized odds ratios other than LOR. Furthermore, in an information-theoretic context, they are expressed as distance models from a parsimonious reference model (the complete symmetry for QS and the independence for AMs), while they satisfy closeness properties with respect to Kullback–Leibler (KL) divergence. Replacing the KL by ϕ divergence, flexible classes of QS-type models for LOR, AMs for LOR, and AMs for generalized odds ratios were generated. However, special QS-type models that are based on homogeneous AMs for LOR have not been extended to ϕ -divergence-based classes so far, or the QS-type models for generalized odds ratios. In this work, we develop these missing extensions, and discuss QS-type models and their generalizations in depth. These flexible families enrich the modeling options, leading to models of better fit and sound interpretation, as illustrated by representative examples.

Keywords: square contingency table; ordinal variables; generalized odds ratios; log-linear models; association models



Citation: Kateri, M. Families of Generalized Quasisymmetry Models: A ϕ -Divergence Approach. *Symmetry* **2021**, *13*, 2297. <https://doi.org/10.3390/sym13122297>

Academic Editor: Kouji Tahata

Received: 26 October 2021

Accepted: 25 November 2021

Published: 2 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A special type of square contingency table that occurs often in studies of correlated or repeated categorical measurements, e.g., in panels or social mobility studies, is a table with row and column classification variables measured on the same scale, which can be nominal or ordinal. Caussinus, in his pioneering work [1], introduced quasisymmetry (QS) models for such square contingency tables, focusing on its mathematical properties and its connection to the models of complete symmetry (S) and marginal homogeneity (MH). The interpretation aspects were developed later by [2–4], among others. The QS model applies on contingency tables that have a common nominal classification scale. An ordinal QS (OQS) model was introduced by [5].

In this work, we focus on the QS model, penetrating its features by considering the alternative equivalent definitions of QS in terms of cell probabilities, local odds ratios (LOR), and as a model measuring departure from the S model. Furthermore, outgoing from the fact that, under QS, the table of LOR is symmetric, we discuss more special parsimonious QS models that are based on the family of association models (AMs) with homogeneous row and column scores ([3,6]). In an information-theoretic setup, QS and OQS models were generalized through ϕ -divergence to the corresponding families of models $QS^{(\phi)}$ and $OQS^{(\phi)}$ ([7,8]). Associated orthogonal decomposition properties of the S model were proved for the $QS^{(\phi)}$ and $OQS^{(\phi)}$ models, respectively, by [7,9]. Further variants of ϕ -divergence QS-type models and S decomposition properties were discussed in [10–12]. A detailed literature review on QS models and links to other models of symmetry and asymmetry can be found in [13]. Here, we introduce a ϕ -divergence family of special QS models that roots on the ϕ -divergence family of AMs ([14]).

Moving from LOR to generalized odds ratios, e.g., global odds ratios (GOR), QS models were also considered for modeling the symmetry of generalized odds ratios ([15]). AMs for generalized odds ratios ([16]) were extended to a broader family through the ϕ -divergence ([17]). Combining these two families of models discussed in [15,17], we introduce here new flexible classes of QS models for generalized odds ratios that are based on ϕ -divergence.

Overall, we revisit the QS model:

- i. Aiming at an indepth discussion of its nature and properties, as consolidated by the alternative possible definitions of QS, and consideration of special QS-type models, with an emphasis on QS-type models that are based on homogeneous AMs.
- ii. Reviewing extensions of QS-type models towards two directions: (a) in an information-theoretic setup by replacing the role of KL divergence by ϕ -divergence, and (b) considering them for generalized odds ratios other than LOR.
- iii. Proposing a new ϕ -divergence family of QS-type models by expanding the relation of QS to homogeneous AMs for the ϕ -divergence AMs.
- iv. Introducing a flexible family of models by extending models of ii(b) above in terms of the ϕ -divergence.

The paper is structured as follows. Section 2 reviews QS and OQS models, focusing on their structural properties, and further discusses AMs with homogeneous scores that are of the QS-type. Section 3 presents QS and OQS models for generalized odds ratios. Generalized families of QS and OQS models for LOR based on ϕ divergence are briefly reviewed in Section 4. In Section 5, the QS-type models of Section 2.2 and the QS models for generalized odds ratios of Section 3 are expanded to corresponding ϕ divergence-based classes of models by modeling ϕ -scaled generalized odds ratios. A selection of the discussed models is illustrated on two examples in Section 6. Section 7 discusses further possible models that can be investigated. Section 8 summarizes our results.

2. Quasisymmetry Models for Square Contingency Tables

Consider an $I \times I$ contingency table cross-classifying two categorical variables X and Y , measured on the same scale, and corresponding to the rows and columns of the table. Let $\pi = (\pi_{ij})$ be the associated probability table with cell entry probabilities $\pi_{ij} = P(X = i, Y = j)$, for $i, j = 1, \dots, I$, where $\pi_{ij} \in [0, 1]$ and $\sum_{i,j} \pi_{ij} = 1$. Then, the QS model, initially introduced by Caussinus [1], is expressed in log-linear form as

$$\log(\pi_{ij}) = u + u_i^X + u_j^X + \alpha_j + u_{ij}^{XY}, \quad i, j = 1, \dots, I, \quad (1)$$

with symmetric interactions, i.e., with

$$u_{ij}^{XY} = u_{ji}^{XY}, \quad i < j, \quad i, j = 1, \dots, I. \quad (2)$$

Parameters in (1) satisfy some identifiability constraints. We set

$$u_1^X = \alpha_1 = 0 \quad \text{and} \quad u_{1i}^{XY} = u_{i1}^{XY} = 0, \quad i = 1, \dots, I. \quad (3)$$

The degrees of freedom of (1) equal $df(QS) = (I - 1)(I - 2)/2$. The α_j parameters measure the departure from model of complete symmetry S , under which $\pi_{ij} = \pi_{ji} = \pi_{ij}^S$, for all $i < j$. Indeed, QS is reduced to S if $\alpha_j = 0$, for all j . Model (1) fits the diagonal entries exactly.

In terms of cell probabilities, (1) is equivalently expressed as

$$\pi_{ij} = \pi_{ij}^S \frac{2c_i}{c_i + c_j}, \quad i, j = 1, \dots, I, \quad (4)$$

with parameters c_i providing insight into sources of marginal inhomogeneity, as discussed in [7,18].

Alternatively, the QS model can be defined in terms of the LOR

$$\theta_{ij}^L = \frac{\pi_{ij}\pi_{i+1,j+1}}{\pi_{i+1,j}\pi_{i,j+1}}, \quad i, j = 1, \dots, I-1, \quad (5)$$

i.e., the odds ratios of all 2×2 subtables formed by pairs of successive rows and successive columns. These form an $(I-1) \times (I-1)$ contingency table, and QS is the model that has symmetric LOR

$$\theta_{ij}^L = \theta_{ji}^L, \quad i < j, \quad i, j = 1, \dots, I-1, \quad (6)$$

and fits the diagonal entries of the probability table exactly, as indicated in [19]. This definition through the LOR highlights a basic structural property of QS, facilitating its physical interpretation and enabling generalizations in new directions by considering alternative types of odds ratios, as we show in Section 3.

2.1. QS Model for Ordinal Classification Variables

Usually in applications of the QS model, the classification scale is ordinal. However, QS is also applicable for tables with nominal classification variables, since it fulfils the permutation invariance property ([20]). In the case of an ordinal scale, alternative QS-type models are possible that are more parsimonious and provide insightful interpretation. Agresti [5] introduced the *ordinal QS (OQS)* model

$$\log(\pi_{ij}) = u + u_i^X + u_j^X + \beta \cdot j + u_{ij}^{XY}, \quad i, j = 1, \dots, I, \quad (7)$$

with interaction parameters satisfying (2) and under (3). In other words, OQS is a special, more parsimonious QS model, derived from (1) when $\alpha_j = \beta \cdot j$. It has just one parameter more than the S model; hence, $df(OQS) = df(S) - 1 = I(I-1)/2 - 1$. Equivalently, (7) can be expressed as

$$\pi_{ij} = \pi_{ji}\delta^{i-j}, \quad i \leq j, \quad i, j = 1, \dots, I, \quad (8)$$

with $\beta = -\log(\delta)$, or as a departure from symmetry model

$$\pi_{ij} = \pi_{ij}^S \frac{\delta^i}{\delta^i + \delta^j}, \quad i \leq j, \quad i, j = 1, \dots, I. \quad (9)$$

Under the OQS model, scores are assigned to the classification categories that equal the category indices ($\mu_j = j, j = 1, \dots, I$). It can be easily verified that these scores in (7) and (8) can equivalently be replaced by any equally spaced scores, i.e., linear transformation of μ_j s. More generally, one could consider an OQS model for known but unequally spaced scores, that is, for any set of known scores $\mu_1 \leq \mu_2 \leq \dots \leq \mu_I$ (with $\mu_1 < \mu_I$). This model however will no more be equivalent to (7). Analogously to the QS model, OQS reduces to the S model when $\beta = 0$ (or $\delta = 1$).

2.2. Association Models with Homogeneous Scores

Log-linear models with interactions for two-way contingency tables are saturated. In case of a square table, the saturated log-linear model is given by

$$\log(\pi_{ij}) = u + u_i^X + u_j^Y + u_{ij}^{XY}, \quad i, j = 1, \dots, I. \quad (10)$$

Association models (AMs) impose special structures on the interactions, thus leading to nonsaturated dependence models of sound interpretation. They are also known as Goodman's AMs (see [6] and references therein). For a detailed discussion of AMs and the associated literature, we refer to [21] (Chapter 6); a short presentation is provided in [22].

The simplest association model is that of *uniform association (U)* that is applied on tables with ordinal classification variables and model interactions through just one parameter of intrinsic association on the basis of equidistant scores assigned to the categories of the

classification variables. The U model for square contingency tables (with classification variables of common scale) can be expressed as

$$\log(\pi_{ij}) = u + u_i^X + u_j^Y + \zeta \mu_i \mu_j, \quad i, j = 1, \dots, I, \quad (11)$$

where ζ is the intrinsic association parameter and $\mu_1 < \mu_2 \dots < \mu_I$ are known scores assigned to the classification categories, which are homogeneous for rows and columns, and equidistant for successive categories ($\mu_{i+1} - \mu_i = \delta > 0$). Under (11), interaction parameters $u_{ij}^{XY} = \zeta \mu_i \mu_j$ are obviously symmetric. Under the U model, all LORs are equal since

$$\log(\theta_{ij}^L) = \zeta(\mu_{i+1} - \mu_i)(\mu_{j+1} - \mu_j) = \zeta \delta^2, \quad i, j = 1, \dots, I-1, \quad (12)$$

and thus (6) is trivially fulfilled. However, (11) is not of the QS-type since it does not exactly fit on the diagonal. Its extension

$$\log(\pi_{ij}) = u + u_i^X + u_j^Y + d_i I(i=j) + \zeta \mu_i \mu_j, \quad i, j = 1, \dots, I, \quad (13)$$

where I is the indicator function, is the homogeneous uniform association model with exactly fitted diagonal entries, denoted by U_{hd} , introduced by [3] as the uniform with main diagonal deleted model. It is a quasisymmetric model for ordinal classifications, more parsimonious than QS with $df(U_{hd}) = (I-1)^2 - I - 1$. It is an alternative to OQS and more parsimonious than OQS for $I > 4$.

Analogously to the OQS model, Model (13) can be considered for arbitrary known scores $\mu_1 \leq \mu_2 \dots \leq \mu_I$ (with $\mu_1 < \mu_I$). Furthermore, considering expression (13) with unknown parametric scores, not necessarily ordered, the homogeneous RC model with exactly fitted diagonals (RC_h^d) is derived, which is another QS-type model, less parsimonious than U_{hd} that can apply also to nominal classification variables. For a discussion on U_{hd} , RC_h^d and further homogeneous AMs of higher order (i.e., homogeneous RC(K) models) and their links to QS, we refer e.g., to [21] (Section 9.4).

3. Generalized QS Models

In case one or both of the classification variables are ordinal, there exist other types of odds ratios that are alternatives to LOR. In our framework of square tables with classification variables measured on the same scale, of interest are, beyond LOR, odds ratios for tables with ordinal classification scale. The most popular type for ordinal classification variables is the global odds ratios (GOR), which for an $I \times I$ contingency table are defined as

$$\theta_{ij}^G = \frac{\pi_{i,j}/\pi_{i,\bar{j}}}{\pi_{\bar{i},j}/\pi_{\bar{i},\bar{j}}} = \frac{(\sum_{s \leq i, k \leq j} \pi_{sk}) / (\sum_{s \leq i, k > j} \pi_{sk})}{(\sum_{s > i, k \leq j} \pi_{sk}) / (\sum_{s > i, k > j} \pi_{sk})}, \quad i, j = 1, \dots, I-1. \quad (14)$$

The characterization *global* is because every θ_{ij}^G is based on the whole contingency table since it dichotomizes X and Y at levels i and j , respectively, and accordingly merges the cell probabilities. GOR treat both classification variables in a symmetric manner. When merging is considered only for one classification variable, for example, Y , while the other is treated locally, then the *cumulative* odds ratios (COR) are derived

$$\theta_{ij}^{CY} = \frac{\pi_{i,j}/\pi_{i,\bar{j}}}{\pi_{i+1,j}/\pi_{i+1,\bar{j}}} = \frac{(\sum_{k \leq j} \pi_{ik}) / (\sum_{k > j} \pi_{ik})}{(\sum_{k \leq j} \pi_{i+1,k}) / (\sum_{k > j} \pi_{i+1,k})}, \quad i, j = 1, \dots, I-1. \quad (15)$$

COR can be used in problems of modeling the effect of an explanatory variable on a response. In particular, θ_{ij}^{CY} could be considered if Y is the response. Obviously, θ_{ij}^{CX} can be analogously defined. For further types of generalized odds ratios and their detailed study, their inter-relations, and associated positive dependence properties, we refer to [23].

Motivated by the definition of QS through the symmetric LOR, the authors in [15] introduced generalized QS-type models for generalized odds ratios other than the LOR. In this context, the classical QS model that applies on the LOR would be denoted by QS^L , while analogously to (6), the QS property for the GOR is defined by

$$\theta_{ij}^G = \theta_{ji}^G, \quad i < j, \quad i, j = 1, \dots, I-1. \quad (16)$$

and denoted by QS^G . On the other hand, the definition of the QS^C model for the COR requires to change also the role of the response variable

$$\theta_{ij}^{C_Y} = \theta_{ji}^{C_X}, \quad i < j, \quad i, j = 1, \dots, I-1, \quad (17)$$

as explained in [15].

4. ϕ -Divergence-Based Families of Generalized QS Models

Model QS and OQS can be defined as departure from complete symmetry models (see (4) and (9)). From a statistical information point of view, they share a common property. Both, under certain conditions (different for each model), are the closest models to complete symmetry when the distance is measured in terms of the KL divergence, as proved by [7,8] for QS and OQS, respectively. Furthermore, the authors in [7,8] introduced and studied general classes of QS and OQS models, derived by replacing the KL divergence by a family of divergences, ϕ -divergence, which includes the KL as special case.

In particular, ϕ -divergence is an important generalized measure for measuring divergence between two probability distributions. In our setup, if $\pi = (\pi_{ij})$, $\mathbf{q} = (q_{ij})$ are two discrete finite $I \times I$ probability distributions, the ϕ -divergence between $\mathbf{q}$ and π is given by

$$I_\phi(\mathbf{q}, \pi) = \sum_{i,j} \pi_{ij} \phi(q_{ij}/\pi_{ij}), \quad (18)$$

where ϕ is a real-valued strictly convex function on $[0, \infty)$ with $\phi(1) = \phi'(1) = 0$, $0\phi(0/0) = 0$, $0\phi(y/0) = \lim_{x \rightarrow \infty} \phi(x)/x$ (see e.g., [24]). The KL divergence

$$I^{KL}(\mathbf{q}, \pi) = \sum_{i,j} q_{ij} \log(q_{ij}/\pi_{ij}), \quad (19)$$

is derived from (18) for $\phi(x) = x \log x$. Setting $\phi(x) = (1-x)^2$, Pearson's divergence is obtained, while $\phi(x) = \frac{x^{\lambda+1}-x}{\lambda(\lambda+1)}$, leads to the power divergence measure of Cressie and Read (CR) [25]

$$I_\lambda^{CR}(\mathbf{q}, \pi) = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^K q_i \left[\left(\frac{q_i}{\pi_i} \right)^\lambda - 1 \right], \quad -\infty < \lambda < \infty, \quad \lambda \neq -1, 0, \quad (20)$$

which is a flexible parametric family itself, controlled by the parameter λ . For $\lambda \rightarrow 0$, (20) converges to the $I^{KL}(\mathbf{q}, \pi)$, for $\lambda = 1$ it corresponds to Pearson's divergence and for $\lambda = -1/2$ to the Hellinger divergence.

The ϕ -divergence-based QS family of models ($QS^{L(\phi)}$), introduced by [7], is defined by

$$\pi_{ij} = \pi_{ij}^S F^{-1}(\alpha_i + \gamma_{ij}), \quad i, j = 1, \dots, I, \quad (21)$$

with $\gamma_{ij} = \gamma_{ji}$, where F^{-1} is the inverse function of $F = \phi'$. For the KL divergence $\phi(x) = x \log(x)$ and (21) leads to model (4) while for the CR $\phi(x) = \frac{x^{\lambda+1}-x}{\lambda(\lambda+1)}$, the $QS^{L(\lambda)}$ family of models

$$\pi_{ij} = \pi_{ij}^S \{1 + \lambda(\alpha_i + \gamma_{ij})\}^{1/\lambda}, \quad i, j = 1, \dots, I, \quad (22)$$

is derived. In this context, the standard QS model is denoted by $QS^{L(0)}$ as a member of this family. The superscript L in these models notation indicates the fact that they impose symmetric restrictions on LOR.

Analogously, the generalized family of OQS models, $OQS^{(\phi)}$, is defined by

$$\pi_{ij} = \pi_{ij}^S F^{-1}(\alpha\mu_i + \gamma_{ij}), \quad i, j = 1, \dots, I, \quad (23)$$

with $\gamma_{ij} = \gamma_{ji}$ and $\mu_1, \dots, \mu_I$ known monotonic equidistant scores (see [8]). $OQS^{(\phi)}$ reduces to (9) for $\phi(x) = x \log(x)$, while, for the CR divergence, $OQS^{(\lambda)}$ is given by

$$\pi_{ij} = \pi_{ij}^S \{1 + \lambda(\alpha\mu_i + \gamma_{ij})\}^{1/\lambda}, \quad i, j = 1, \dots, I. \quad (24)$$

5. New Families of ϕ -Divergence Generalized QS Models

The QS-type models of Section 2.2 that are linked to AMs with homogenous row and column scores can be extended to ϕ -divergence-based families through the ϕ -divergence AMs of [14]. A brief presentation of the ϕ -divergence AMs and the underlying concept can be found in [22]. Here, we focus just on the families corresponding to the U model with homogeneous row and column scores. Associated ϕ -divergence-based family of models $U_h^{(\phi)}$ is defined as

$$\pi_{ij} = \pi_{i+}\pi_{+j}F^{-1}(\alpha_i + \beta_j + \zeta\mu_i\mu_j), \quad i, j = 1, \dots, I, \quad (25)$$

where π_{i+} and π_{+j} denote the i -th row and j -th column marginals respectively, i.e., $\pi_{i+} = \sum_{j=1}^I \pi_{ij}$ and $\pi_{+j} = \sum_{i=1}^I \pi_{ij}$. For the KL divergence, $U_h^{(\phi)}$ leads to model (11) while for the CR-divergence it takes the form

$$\pi_{ij} = \pi_{i+}\pi_{+j} \left[\frac{1}{\lambda + 1} + \lambda(\alpha_i + \beta_j + \zeta\mu_i\mu_j) \right]^{1/\lambda}, \quad i, j = 1, \dots, I, \quad (26)$$

denoted by $U_h^{(\lambda)}$. Hence the special QS-type model U_{hd} defined in (13), can be extended to a family of models

$$\pi_{ij} = \pi_{i+}\pi_{+j}F^{-1}(\alpha_i + \beta_j + d_i I(i=j) + \zeta\mu_i\mu_j), \quad i, j = 1, \dots, I, \quad (27)$$

denoted by $U_{hd}^{(\phi)}$. The model expression corresponding to CR divergence is denoted by $U_{hd}^{(\lambda)}$.

Furthermore, the QS models for generalized odds ratios, introduced in [15] and presented in Section 3, can be extended to a flexible ϕ -divergence family, based on the ϕ -divergence generalized AMs of [17], which are briefly presented below, adjusted in our set-up.

Analogously to expression (12) for the U_h^L model, $U_h^{(\phi)}$ can alternatively be expressed as

$$\theta_{ij}^{L(\phi)} = \zeta(\mu_{i+1} - \mu_i)(\mu_{j+1} - \mu_j) = \zeta\delta^2, \quad i, j = 1, \dots, I-1, \quad (28)$$

where

$$\theta_{ij}^{L(\phi)} = F\left(\frac{\pi_{ij}}{\pi_{i+}\pi_{+j}}\right) + F\left(\frac{\pi_{i+1,j+1}}{\pi_{i+1,+}\pi_{+,j+1}}\right) - F\left(\frac{\pi_{i+1,j}}{\pi_{i+1,+}\pi_{+j}}\right) - F\left(\frac{\pi_{i,j+1}}{\pi_{i+}\pi_{+,j+1}}\right), \quad (29)$$

for $i, j = 1, \dots, I-1$, are measures of local dependence, scaled through the ϕ -divergence and denoted by $LOR^{(\phi)}$. For $\phi(x) = x \log x$, $LOR^{(\phi)}$ is the log(LOR), modeled in (12), which in the sequel is denoted as $\theta_{ij}^{L(0)} = \log(\theta_{ij}^L)$, $i, j = 1, \dots, I-1$. For the CR divergence, (29) becomes

$$\theta_{ij}^{L(\lambda)} = \frac{1}{\lambda} \left[\left(\frac{\pi_{ij}}{\pi_{i+}\pi_{+j}} \right)^\lambda + \left(\frac{\pi_{i+1,j+1}}{\pi_{i+1,+}\pi_{+,j+1}} \right)^\lambda - \left(\frac{\pi_{i+1,j}}{\pi_{i+1,+}\pi_{+j}} \right)^\lambda - \left(\frac{\pi_{i,j+1}}{\pi_{i+}\pi_{+,j+1}} \right)^\lambda \right]. \quad (30)$$

Forcina and Kateri ([17]) provided expressions for ϕ -scaled generalized odds ratios and introduced families of ϕ -divergence AMs for generalized odds ratios, which they studied. Thus, for example, GOR and COR extend to $\text{GOR}^{(\phi)}$ and $\text{COR}^{(\phi)}$, given by

$$\theta_{ij}^{G(\phi)} = F\left(\frac{\pi_{i,j}}{\pi_{i+}\pi_{+j}}\right) + F\left(\frac{\pi_{\bar{i},\bar{j}}}{\pi_{\bar{i}+}\pi_{+\bar{j}}}\right) - F\left(\frac{\pi_{\bar{i},j}}{\pi_{\bar{i}+}\pi_{+j}}\right) - F\left(\frac{\pi_{i,\bar{j}}}{\pi_{i+}\pi_{+\bar{j}}}\right), \quad (31)$$

and

$$\theta_{ij}^{C(\phi)} = F\left(\frac{\pi_{i,j}}{\pi_{i+}\pi_{+j}}\right) + F\left(\frac{\pi_{i+1,\bar{j}}}{\pi_{i+1,+}\pi_{+\bar{j}}}\right) - F\left(\frac{\pi_{i+1,j}}{\pi_{i+1,+}\pi_{+j}}\right) - F\left(\frac{\pi_{i,\bar{j}}}{\pi_{i+}\pi_{+\bar{j}}}\right), \quad (32)$$

for $i, j = 1, \dots, I - 1$. The merged probabilities in (31) and (32) are the same as defined in (14) and (15). For the KL divergence, (31) and (32) reduce to (14) and (15), while for the CR divergence, setting $F(x) = \frac{1}{\lambda} x^\lambda$, the corresponding expressions $\text{GOR}^{(\lambda)}$ and $\text{COR}^{(\lambda)}$ are derived. Through these ϕ -scaled generalized odds ratios, generalized QS models, such as QS^G and QS^C , are extended to ϕ -divergence-based families of models by replacing in the definitions (16) and (17) the GOR and COR by $\text{GOR}^{(\phi)}$ and $\text{COR}^{(\phi)}$. QS-type models for other types of generalized odds ratios, introduced in [15], can be analogously extended to ϕ -divergence-based families.

6. Examples

We first illustrate QS-type models on one of the most classical datasets of square tables, namely, the women vision data provided in Table 1. Apart from the standard QS model fitted often in the literature, the author in [5] fitted on this dataset the OQS model, while the authors in [7] the $QS^{L(\lambda)}$ and in [8] the $OQS^{(\lambda)}$, for $\lambda = 1$. Furthermore, the authors in [15] fitted QS models for generalized odds ratios other than the LOR.

Table 1. Cross-classification of 7477 women by unaided distance vision of right and left eyes. In parentheses are the MLEs of the expected frequencies under $OQS^{(1)}$ model.

Right Eye Grade	Left Eye Grade				Total
	Best	Second	Third	Worst	
best	1520	266 (263.35)	124 (133.37)	66 (59.17)	1976
second	234 (236.65)	1512	432 (418.20)	78 (88.54)	2256
third	117 (107.63)	362 (375.80)	1772	205 (202.25)	2456
worst	36 (42.83)	82 (71.46)	179 (181.74)	492	789
total	1907	2222	2507	841	7477

Our second example, provided in Table 2, cross-classifies male respondents of the 2008 General Social Survey (GSS) in the USA on the basis of their degree of pride with regard to America's economic vs scientific and tech achievements. We applied on this dataset the same models as on the women-vision data.

Table 2. Cross-classification of male respondents of GSS 2008 survey by degree of proud (VP: very proud, SWP: somewhat proud, NVP: not very proud; NP: not proud at all) with regard to America's economic achievements (rows) vs scientific and tech achievements (columns). In parentheses are the MLEs of the expected frequencies under $QS^{G(0)}$ model.

Economic	Science and Tech				Total
	VP	SWP	NVP	NP	
VP	369 (368.81)	59 (58.73)	6 (6.84)	1 (0.97)	435
SWP	226 (226.63)	238 (237.69)	10 (9.74)	3 (2.85)	477
NVP	60 (59.16)	67 (67.65)	14 (13.78)	2 (2.09)	143
NP	7 (7.063)	16 (16.18)	3 (2.83)	2 (2.00)	28
total	662	380	33	8	1083

In Table 3 we provide the likelihood ratio goodness of fit (GOF) test statistic values for QS-type models fitted on the LOR for both examples. Table 4 shows the GOF test statistics for the generalized QS model fitted on the GOR.

Table 3. Goodness-of-fit test statistics (with corresponding P values in parentheses) for the $QS^{L(\lambda)}$ and $OQS^{(\lambda)}$ for $\lambda \in \{0, 1\}$ along with the U_{hd} model fitted on data of Tables 1 and 2.

Model	df	Table 1	Table 2
$QS^{L(0)}$	3	7.2708 (0.0638)	3.4181 (0.3315)
$QS^{L(1)}$	3	7.2620 (0.0640)	20.1060 (0.0002)
$OQS^{(0)}$	5	7.2804 (0.2006)	8.8867 (0.1137)
$OQS^{(1)}$	5	7.2708 (0.2013)	47.7122 ($<10^{-5}$)
U_{hd}	4	61.3082 ($<10^{-5}$)	3.9002 (0.4197)

For the women-vision example, OQS models were best. The scale in this case is not of practical importance, since $QS^{L(0)}$ and $QS^{L(1)}$, corresponding to the KL and Pearsonian divergence, are of comparable fit (their maximum likelihood estimates (MLEs) differ on the second decimal place). The MLEs of the expected cell frequencies under $QS^{L(1)}$ are provided in Table 1. This model, i.e., (24) with $\lambda = 1$ and $\mu_i = i, i = 1, \dots, I$, takes the final form (see [8])

$$\pi_{ij} = \pi_{ij}^S \{1 + \alpha(i - j)\}, \quad i, j = 1, \dots, I. \quad (33)$$

The MLE for α is $\hat{\alpha} = -0.0534 < 0$ and hence the probabilities in the lower triangle in Table 1 are estimated to be smaller than those in the upper. Hence the vision is worse for the left eye. Under this model, it holds

$$\frac{\pi_{ij}}{\pi_{ji}} = \frac{1 + \alpha(i - j)}{1 + \alpha(j - i)}, \quad i < j.$$

Thus, the odds of an observation falling in a certain subdiagonal under the main diagonal of the table (instead of the corresponding superdiagonal) are estimated as $\frac{\hat{\pi}_{ij}}{\hat{\pi}_{ji}} = \frac{1 - 0.0534(i - j)}{1 - 0.0534(j - i)}, i < j$. Notice that in [8] the corresponding $\hat{\alpha}$ value is different ($=0.119$). This is due to rescaling, since there is used a different set of u_i scores (the $\hat{\pi}_{ij}$'s are the same).

The situation is different for the second example, where it is clear that the KL divergence should be used for modeling the LOR (see Table 3). Furthermore we see that for this data set, the U_{hd} model, that imposes a special parsimonious structure on the interaction terms and not on the main effects (as under the OQS models), is of better fit. However, an impressive fit is provided by the QS^G models (see Table 4). The best fit is for $\lambda = 0$ and thus the MLEs of the expected cell frequencies under $QS^{G(0)}$ are shown in Table 2. Hence, for this data set the QS property is significantly stronger supported for the global (than the local) dependencies.

Table 4. Goodness-of-fit test statistics for the $QS^{G(\lambda)}$ models for the GOR, fitted on the data of Tables 1 and 2 for $\lambda \in \{-1/2, 0, 1/3, 2/3, 1\}$. All models have $df = 3$.

Model	Table 1	Table 2
$QS^{G(-1/2)}$	7.5474 (0.0564)	0.4239 (0.9353)
$QS^{G(0)}$	7.6242 (0.0544)	0.1652 (0.9830)
$QS^{G(1/3)}$	7.8872 (0.0484)	0.2105 (0.9759)
$QS^{G(2/3)}$	8.4732 (0.0372)	0.4437 (0.9311)
$QS^{G(1)}$	9.2519 (0.0261)	0.8973 (0.8261)

In our examples, we considered specific choices for the parameter λ . Analysis and interpretation of results follows analogously for other choices of the parameter λ .

7. Discussion

In future research, it would be interesting to consider more parsimonious QS -type models for generalized odds ratios, analogs to U_{hd} for LOR, as for example, the model of uniform GOR with homogeneous (equidistant) row and column scores, i.e., satisfying

$$\log(\theta_{ij}^G) = \zeta \delta^2, i, j = 1, \dots, I - 1,$$

that additionally fits the probabilities on the main diagonal cells exactly. AMs that model interactions other than local, though they are naturally defined by the corresponding type of odds ratios, do not provide closed form expressions for the individual cell probabilities. Hence, a definition of the associated QS -type models by expressions analog to (13) is not possible. For the same reason, the OQS model cannot be extended to other types of odds ratios by the approach adopted here. Since it imposes a special structure on the main effects, this cannot be captured when defining models in terms of odds ratios; expressions in terms of cell probabilities are required. Recently, the authors in [17] derived expressions for such generalized AMs in terms of suitable associated marginal probabilities. These expressions include parameters for the main effects (see Forms (9) and (10) in [17]). One could generalize U_{hd} and OQS for other types of odds ratios, adopting the framework of [17].

8. Conclusions

In this work, we revisited the QS model for square contingency tables with commensurable classification variables and discussed its possible equivalent formulations. QS is mostly expressed in terms of cell probabilities, while it can alternatively be expressed in terms of local odds ratios. Its definition as a departure model from the more parsimonious model of complete symmetry provides additional interpretation features. Furthermore, we considered the OQS model, a more parsimonious QS -type model, applicable if the classification scale is ordinal, which imposes a special structure on the main effects of the model. On the other hand, further QS -type models can be derived by considering a special structure for the interaction terms. This is possible through AMs with homogeneous row and column scores. In particular, by adding parameters to homogeneous AMs that ensure the exact fit on the diagonal entries of the contingency table, models are derived that model the off-diagonal cells and have symmetric interaction terms. Thus, they are of the QS -type, but more parsimonious than the standard QS model. All these models are related to LOR and model local dependencies of the table. Next we present how these models can be defined for other types of generalized odds ratios, reviewing the work of [15].

In a statistical information-theoretic setup, QS -type models and AMs satisfy properties of closeness to a specific reference model, when their divergence from the reference model is measured in terms of KL divergence. The reference model is that of symmetry (for QS and OQS models) or independence (for AMs). Replacing KL divergence with ϕ -divergence, generalized families of AMs, QS and OQS models for LOR were considered by [7,8,14], respectively. The QS model was linked to GOR in [16], while [15] introduced

and implemented QS models for GOR and other types of generalized odds ratios, without, however, considering the link to divergence measures and associated properties. The possible extension of these models in terms of ϕ -divergence was a topic for further research in [15]. Here, we extended these models to ϕ -scaled generalized odds ratios and linked them to corresponding AMs on the basis of the results and models discussed in [17]. We demonstrated the flexibility in modeling the classes of models discussed here by implementing and discussing some of these models on two representative examples.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data of Table 2 can be found in <https://gssdataexplorer.norc.ox.ac.uk/> (accessed on 10 October 2021).

Acknowledgments: The author thanks the reviewers for their useful comments on an earlier version of the manuscript.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Caussinus, H. Contribution à l'analyse statistique des tableaux de corrélation. In *Annales de la Faculté des Sciences de Toulouse; Université Paul Sabatier: Toulouse, France*, 1965; Volume 29, pp. 77–183.
2. Bishop, Y.M.; Fienberg, S.E.; Holland, P.W. *Discrete Multivariate Analysis: Theory and Practice*; MIT Press: Cambridge, MA, USA, 1975.
3. Goodman, L.A. Simple Models for the Analysis of Association in Cross-Classifications having Ordered Categories. *J. Am. Stat. Assoc.* **1979**, *74*, 537–552. [\[CrossRef\]](#)
4. McCullagh, P. Some applications of quasisymmetry. *Biometrika* **1982**, *69*, 303–308. [\[CrossRef\]](#)
5. Agresti, A. A simple diagonals–parameters symmetry and quasi-symmetry model. *Stat. Probab. Lett.* **1983**, *1*, 313–316. [\[CrossRef\]](#)
6. Goodman, L.A. The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Ann. Stat.* **1985**, *13*, 10–69. [\[CrossRef\]](#)
7. Kateri, M.; Papaioannou, T. Asymmetry models for contingency tables. *J. Am. Stat. Assoc.* **1997**, *92*, 1124–1131. [\[CrossRef\]](#)
8. Kateri, M.; Agresti, A. A class of ordinal quasi symmetry models for square contingency tables. *Stat. Probab. Lett.* **2007**, *77*, 598–603. [\[CrossRef\]](#)
9. Saigusa, Y.; Tahata, K.; Tomizawa, S. Orthogonal decomposition of symmetry model using the ordinal quasi-symmetry model based on f -divergence for square contingency tables. *Stat. Probab. Lett.* **2015**, *101*, 33–37. [\[CrossRef\]](#)
10. Tahata, K. On asymmetry models and decompositions of symmetry for square contingency tables with ordered categories. In *Proceedings of the 32nd International Workshop on Statistical Modelling, Groningen, The Netherlands, 3–7 July 2017; Volume 2*, pp. 231–234.
11. Tahata, K. Separation of symmetry for square tables with ordinal categorical data. *Jpn. J. Stat. Data Sci.* **2020**, *3*, 469–484. [\[CrossRef\]](#)
12. Fujisawa, K.; Tahata, K. Asymmetry model based on f -divergence and orthogonal decomposition of symmetry for square contingency tables with ordinal categories. *SUT J. Math.* **2020**, *56*, 39–53.
13. Tahata, K.; Tomizawa, S. Symmetry and asymmetry models and decompositions of models for contingency tables. *SUT J. Math.* **2014**, *50*, 131–165.
14. Kateri, M.; Papaioannou, T. f -divergence Association Models. *Int. J. Math. Stat. Sci.* **1994**, *3*, 179–203.
15. Kateri, M.; Gottard, A.; Tarantola, C. Generalized quasi symmetry models for ordinal contingency tables. *Aust. N. Z. J. Stat.* **2017**, *59*, 239–253. [\[CrossRef\]](#)
16. Bartolucci, F.; Forcina, A. Extended RC association models allowing for order restrictions and marginal modeling. *J. Am. Stat. Assoc.* **2002**, *97*, 1192–1199. [\[CrossRef\]](#)
17. Forcina, A.; Kateri, M. A new general class of RC association models: Estimation and main properties. *J. Multivar. Anal.* **2021**, *184*, 104741. [\[CrossRef\]](#)
18. Becker, M.P. Quasisymmetric models for the analysis of square contingency tables. *J. R. Stat. Soc. Ser. (Methodol.)* **1990**, *52*, 369–378. [\[CrossRef\]](#)
19. Goodman, L.A. Multiplicative models for square contingency tables with ordered categories. *Biometrika* **1979**, *66*, 413–418. [\[CrossRef\]](#)
20. McCullagh, P. A class of parametric models for the analysis of square contingency tables with ordered categories. *Biometrika* **1978**, *65*, 413–418. [\[CrossRef\]](#)
21. Kateri, M. *Contingency Table Analysis: Methods and Implementation Using R*; Birkhäuser (Springer): New York, NY, USA, 2014.
22. Kateri, M. ϕ -divergence in contingency table analysis. *Entropy* **2018**, *20*, 324. [\[CrossRef\]](#) [\[PubMed\]](#)

-
23. Douglas, R.; Fienberg, S.E.; Lee, M.L.T.; Sampson, A.R.; Whitaker, L.R. Positive dependence concepts for ordinal contingency tables. In *Topics in Statistical Dependence*; Lecture Notes–Monograph Series; Block, H.W., Sampson, A.R., Savits, T.H., Eds.; Institute of Mathematical Statistics: Ann Arbor, MI, USA, 1990; pp. 189–202.
 24. Pardo, L. *Statistical Inference Based on Divergence Measures*; Chapman & Hall: New York, NY, USA, 2006.
 25. Cressie, N.; Read, T. Multinomial Goodness-of-Fit Tests. *J. R. Stat. Soc. Ser. B* **1984**, *46*, 440–464. [[CrossRef](#)]