University of Toronto
Department of Economics

Working Paper 757

A Coefficient of Variation for Multivariate Ordered Categorical Outcomes.

By Gordon Anderson

September 05, 2023

A Coefficient of Variation for Multivariate Ordered Categorical Outcomes.

Gordon Anderson.

Economics Department University of Toronto.

June 2023

Summary.

Comparing the relative variation of ordinal variates defined on diverse populations is challenging. Pearsons' Coefficient of Variation or its inverse (the Sharpe Ratio), each used extensively for comparing relative variation or risk tempered location in cardinal paradigms, cannot be employed in ordinal data environments unless cardinal scale is attributed to ordered categories. Unfortunately, due to the scale dependencies of the Coefficient of Variations denominator and numerator, such arbitrary attribution can result in equivocal comparisons. Here, based upon the notion of probabilistic distance, unequivocal, scale independent, Coefficient of Variation and Sharpe Ratio analogues for use with Multivariate Ordered Categorical Data are introduced and exemplified in an analysis of Canadian Human Resource distributions.

Introduction.

Is the diversity of a society's political opinions similar to its diversity of views about the economy? Is the variation in self-reported happiness responses of a group matched by the variation in their responses to self-reported ill health? How can the diversity of treatment responses across treatment groups be compared when those responses are ordinal in nature? Is the variation in responses to a given categorical questionnaire common or dissimilar across respondent types? Does a group exhibit the same diversity in its response to distinctly different questionnaires? All such questions are posed in the generic context of analyzing and comparing the variability of ordered categorical responses of diverse groups. Usually, in cardinally measurable paradigms, the Coefficient of Variation (COV) has been used to answer the generic question but cardinality is of the essence and the equivocation inherent in the attribution of an arbitrary cardinal scale to ordinal outcomes (different, equally valid scales can yield substantively different conclusions) invalidates it as a solution. Thus, the ever-increasing use of Ordered Categorical Data in all spheres of the social and physical sciences presents a measurement challenge, particularly when analyzing between group differences of within group variability.

First introduced by Pearson (1896) as the ratio of the standard deviation to the mean, COV is a unit free relative variation measure. It, or its inverse (the well-known Sharpe (1964,1994) Ratio used for examining risk adjusted Excess Returns[1]), have been used extensively in economics and finance as a measure of economic inequality and relative risk. Despite its disadvantages (Kvalseth 2017), it has also seen extensive use in the physical and biological sciences (Weber et. al 2004), engineering (Jalalibal et.al. 2021) and Industrial Organization fields (Bedeian & Mossholder 2000) where cardinally measurable data abounds. Pearson proposed COV in response to Galtons' practice of using the 13 to 12 male-female size ratio[2] in his work on Natural Inheritance (Galton 1894) and used the mean focused variation measure standardized by the mean to address the comparative variation of organ sizes (usually skull and bone dimensions) across race and gender[3]. His rationale for standardizing the dispersion measure by the

---

[1] Following concerns that the standard deviation was not an adequate reflection of downside risk the Sortino Ratio (Sortino and Price 1994) modified the Sharpe Ratio by considering only the non-positive deviations from the mean in the standard deviation calculation.

[2] Galton would rescale a female organ size by 13/12 to obtain a comparable male equivalent.

[3] Pearson's view of his new statistic (Pearson 1896 pp. 276-9) was circumspect but enthusiastic, he wrote: "*Of course, it does not follow because we have defined in this manner our "coefficient of variation", that this is really a significant quantity in the comparison of various races; it may be only a convenient mathematical*

mean was a concern for reliability and consistency across disparate distributions, that measurement should not be too variable or at least consistently variable i.e. sufficiently stable about the mean value, so as to be comparably useful across races and genders.

In more recent times COV has seen a variety of extensions to multivariate environments (see Albert and Zhang 2010 for a survey) based on alternative approaches to dealing with the multiplicity of measurement units. Clearly cardinality is of the essence in establishing a mean value as well as variation about. Thus COV has no natural analogue in multivariate ordered categorical data environments since ordered categorical data are bereft of cardinal measure, unless it is arbitrarily endowed, but that presents problems of scale dependency and concomitant equivocation (Bond and Lang 2019, Schroder and Yitzhaki 2017) which raises questions about the viability of measurement of variation about a location measure in ordered categorical situations.

In the absence of cardinal measure, researchers have used notions of probabilistic distance (Mendelson 1987) and the construct of a median preserving spread in order to quantify variation for the purpose of measuring inequality and polarization (Blair and Lacy 2000, Allison and Foster 2004, Kobus and Kurek 2019). The probabilistic distance of a given category from the median focus category is measured in terms of the likelihood of an outcome in the given or any other category between it and the median category, the higher that probability is, the further apart are the categories deemed to be. Inequality is then quantified as the average probabilistic distance from the median focus category of all non-median categories.

If inequality is conceptualized as the antithesis of complete equality, the aggregate distance of subjects from a potential focus point of complete commonality would characterize it. In this context the Median or the Mean may not be very good focus points since they can often have low probabilistic density (for example in heavily skewed or strongly segmented bimodal distributions) which renders them unlikely candidates as point of complete equality and hence, given this antithetic view of inequality, less useful as focus points. Furthermore, in multivariate settings mean and median focal points are difficult to determine whereas the modal point is usually uniquely determined (even in a multiplicity of nodes there is usually one node with a density greater than the rest). Noting that, in a

---

expression, but I believe there is evidence to show that it is a more reliable test of "efficiency"[3] in a race than absolute variation."

likelihood sense, the mode is the most likely point of complete equality, Anderson and Yalonetzki (2023) provide an alternative to the Median Preserving Spread formulation in the Modally Preserving Spread with the mode as a focal point. As an inequality measure it has a natural likelihood-based interpretation (average probabilistic distance from the most likely point of commonality), is well defined in multidimensional situations, and has a probabilistic unit of measurement which is common to all dimensions. Pearsons' concern in ordered categorical data environments would have been that such measures of spread would not be "stable" for reliable comparison across groups without suitable locational standardisation. Sharpes' concern would have been that the location measure would have not been diluted by an appropriate measure of uncertainty in essence it is a variation standardised location measure much like a standard normal statistic. All that remains for an Ordered Categorical Coefficient of Variation or its inverse ($OCCOV$ or $OCCOV^{-1}$) is to render such a measure unit free by standardising the modally focused variation measure with some relevant probability-based measure of location.

In the following, details of the conventional Coefficient of Variation and its multivariate versions are outlined in Section 1, Section 2 proposes an analogue for multivariate ordered categorical data environments and Section 3 provides an exemplifying application to the analysis of Canadian Human Resource Stock variation across gender and time (the 2006-2016 decade). To anticipate the results, it was found that, whereas Female outcome relative variation was reasonably stable over time and initially similar to Male relative variation, Male outcomes exhibited a dramatic shift in both variation and location over the decade. Generally relative variation in education and training was greater than relative variation in experience. Conclusions are drawn in Section 4.

1. The Coefficient of Variation.

Consider a continuous cardinally measurable variable x with $0 < x < Y < \infty$, denote the group t distribution $f_t(x)$ with a corresponding CDF: $F_t(x) = P_t(X < x) = \int_0^x f_t(z)dz$, Survival Function SF: $S_t(x) = P_t(X \geq x) = 1 - F_t(x)$, group t mean as $\mu_t = E_{f_t(x)}(x) = \int_0^Y x f_t(x)dx$ and group t variance of x as $\sigma_t^2 = E_{f_t(x)}((x-\mu_t)^2) = \int_0^Y (x-\mu_t)^2 f_t(x)dx$. Note that integrating the mean formula by parts reveals it to be equivalent to the integral of the survival function[4] so that $\mu_t = \int_0^Y S_t(x)dx$ which yields an alternative interpretation of the mean as the cumulation of chances of higher outcomes than x over the range of x.

Then $COV_{f_t(x)}(x)$, the group t Coefficient of Variation may be written as:

$$COV_{f_t(x)}(x) = \frac{\sqrt{\sigma_t^2}}{\mu_t} = \frac{\sqrt{\sigma_t^2}}{\int_0^Y S_t(x)dx} \qquad [1]$$

Practically, for a collection of N randomly sampled cardinally measurable values $x_i, i = 1,..N$, where for convenience $x_i \geq 0$, the basic COV is given by[5]:

$$COV = \frac{\sqrt{\sum_{i=1}^N (x_i - \underline{x})^2/(N-1)}}{\underline{x}} = \frac{\hat{\sigma}}{\hat{\mu}}; where\ \hat{\mu} = \underline{x} = \sum_{i=1}^N x_i\,/N \qquad [2]$$

As a measure of relative variation, it can be seen to be the square root of the variance estimate, which is the average of the squared distances of the x_i's from the mean, divided by the mean. The mean is very much the focus of the statistic, it is the value that minimises the magnitude of the variance estimate (a similar variation measure around any other value would always be larger) and division by it dilutes the standard deviation value and renders the statistic a unit free measure[6]. Its inverse, known as the Sharpe Ratio (Sharpe 1994), is employed in the finance field in risk vs. return scenarios, where the x_i's are rates of excess return and $\hat{\sigma}$ is a measure of their riskiness level, COV^{-1} can be seen to dilute the average excess return level by the level of riskiness (much like standardizing a normal random variable by its standard deviation to yield a standard normal variate). In the context of income

---

[4] Integration by parts yields: $\int_0^Y u(x)v'(x)dx = [u(x)v(x)]_0^Y - \int_0^Y u'(x)v(x)dx$. Let $u(x) = x$ and $v'(x) = f(x)$, then: $\mu = \int_0^Y xf(x)dx = [xF(x)]_0^Y - \int_0^Y F(x)dx = Y - \int_0^Y F(x)dx = \int_0^Y (1 - F(x))dx = \int_0^Y S(x)dx.$

[5] When data are sampled from a set of K discrete cardinally measurable values $x_k\ k = 1,..,K$ where p_k is the proportion of the sample that took on the value x_k, [1] can be computed as:

$$COV = \frac{\sqrt{\sum_{k=1}^K (x_k - \underline{x})^2 p_k}}{\underline{x}} = \frac{\hat{\sigma}}{\hat{\mu}}; where\ \hat{\mu} = \underline{x} = \sum_{k=1}^K x_k\, p_k \qquad [2a]$$

[6] Similar statistics can be contrived if other foci are of interest by making $\underline{x}$ the median or modal value of the collection, indeed dividing the standard deviation by any quantile value would render it a unit free measure relative to the designated quantile.

inequality and wellbeing measurement[7], the Sharpe measure dilutes the average income level by a measure of the inequality with which incomes are distributed.

Multivariate Extensions of the Coefficient of Variation.

Extending the coefficient of variation to the multivariate paradigm has seen several alternative formulations of a Multivariate COV proposed in the literature. Albert and Zhang (2010) reviewed some of the alternatives and proposed a novel formulation themselves, all amount to standardizing a function of the variance covariance matrix with the inner product of the dimension means and taking the square root thereof. The object being to reconcile the diverse units of measurement in the various dimensions to obtain a unit free measure. In a K >1 dimension setting, letting $\underline{\mu}$ be the Q x 1 vector of dimension means and C be the Q x Q covariance matrix, the alternatives (see inter alia Reyment 1960, Van Valen 1974, Voinov and Nikulin 1966 and Albert and Zhang 2010) considered by Albert and Zhang were:

$$\sqrt{\frac{det(C)^{1/K}}{\underline{\mu}'\,\underline{\mu}}};\quad \sqrt{\frac{trC}{\underline{\mu}'\,\underline{\mu}}};\quad \sqrt{\frac{1}{\underline{\mu}'C^{-1}\,\underline{\mu}}};\quad \sqrt{\frac{\underline{\mu}'\,C\underline{\mu}}{\left(\underline{\mu}'\,\underline{\mu}\right)^{2}}}.$$

It is worthy of note that when Q = 1 all of these formulae reduce to the conventional coefficient of variation yet in the multivariate empirical setting they can yield very different values for a given sample (Aerts, Haesbroeck and Ruwet 2015).

2. A Coefficient of Variation Analogue for Multivariate Ordered Categorical Data.

To develop Coefficient of Variation or a Sharpe-Ratio analogues for Ordered Categorical data, a means of measuring variation in the absence of cardinal measure is required together with a means of standardizing it with an appropriate location measure in that context. The notion of probabilistic distance, the sense that two ordered outcomes are further apart the greater is the probability of an outcome between them occurring, is useful in this case.

[7] COV can be shown to satisfy the inequality measurement axioms of anonymity, scale invariance and population independence (Champernowne and Cowell 1999).

When data are ordered categorical (suppose there to be K such ordered categories) it is not possible to compute a mean or the individual squared distances from it without arbitrary attribution of cardinal value to the respective categories which invites concerns with respect to ambiguity and scaling effects. To circumvent these problems in quantifying inequality, Allison and Foster (2004) resorted to the notion of probabilistic distance first introduced in Mendelson (1987). Contemplate K ordered categories indexed $k = 1,..,K$ with a Probability Distribution Function f described by the probabilities $p_{fk}, k = 1,..,K$ where $\sum_{k=1}^{K} p_{fk} = 1$. For a given outcome $k^* \in 1,..,K$ and outcomes $k = k^* + 1,..,K$, define the Upper Cumulants of f with respect to k^* as $F_k^{U,k^*} = \sum_{i=k^*+1}^{k} p_{fi}$ (note for $k \leq k^*$, $F_k^{U,k^*} = 0$) and, for outcomes $k = 1,..,k^* - 1$, define its Lower Cumulants as $F_k^{L,k^*} = \sum_{i=k}^{k^*-1} p_{fi}$ (note for $k \geq k^*$, $F_k^{L,k^*} = 0$). It may be seen that $\frac{F_k^{L,k^*}}{F_1^{L,k^*}}$ $k = 1,..,k^* - 1$ is in effect the SF of the below k^* conditional PDF, whereas $\frac{F_k^{U,k^*}}{F_K^{U,k^*}}$ $k = k^* + 1,..,K$ is the CDF of the above k^* conditional PDF. When $k > k^*$, F_k^{U,k^*} is the probability of an outcome between k^* and $k + 1$ occurring which is monotonically non decreasing in k, when $k < k^*$, F_k^{L,k^*} is the probability of an outcome between k^* and $k - 1$ occurring which is monotonically non-decreasing in $k^* - k$. Each record a sense of probabilistic distance of k from k^* in terms of the chance that an outcome will emerge between k and k^* which increases with $|k^* - k|$. Similarly defining G_k^{U,k^*}, G_k^{L,k^*}, the Upper and Lower Cumulants of g about k^*, then g constitutes an increasing spread of f with respect to outcome k^* when:

$$G_k^{L,k^*} \geq F_k^{L,k^*} \forall\, k = 1,..,k^* - 1 \; and \; G_k^{U,k^*} \geq F_k^{U,k^*} \forall\, k = k^* + 1,..,K \; with > somewhere. \qquad [3]$$

The Mendelson (1987) condition [3] amounts to a first order stochastic dominance condition on the "downward looking" below k^* conditional distributions (i.e. imagine the category orderings below k^* were reversed) and the "upward looking" above k^* conditional distributions where f dominates g in each context. Intuitively, with respect to k^* inequality in g distribution is greater than inequality in f distribution with respect to k^* when the chance of below k^* outcomes and the chance of above k^* outcomes are both at least as great in g as they are in f with strictly greater than in at least one case[8].

Given the absence of cardinal measure, setting k^* as the "median" category and using this notion of probabilistic distance has been the basis of inequality and bi-polarization measurement in univariate

[8] This construct is similar to notions of left and right distributional separation developed in Anderson (2004).

ordered categorical paradigms (Blair and Lacy 2000, Allison and Foster 2004, Kobus 2015). However, the median outcome could well be an unlikely event and, if inequality is construed as the antithesis of complete commonality or equality in the population, it would not serve as a good focus point for a likelihood-based inequality measure.

The Modal Preserving Spread.

Define the Modal outcome of distribution f as outcome k^* *such that* $p_{fk^*} = \max_k p_{fk}$. Determining k^* by seeking that category for which $\hat{p}_{fk^*} = \max_k \hat{p}_{fk}$ where $\hat{p}_{fk}, k = 1,..,K$ are the maximum likelihood estimates of category densities, renders k^* as the maximum likelihood estimate of the category most likely to command unanimity of membership. Since the smallest possible value of p_{fk^*} is $\frac{1}{K} + \varepsilon$ where ε is an arbitrarily small positive value, $\frac{1}{K} < p_{fk^*} \leq 1$ and, when p_{fk^*} is viewed as the chance that the whole population resides in outcome k^*, $LC(f) = (Kp_{fk^*} - 1)/(K - 1)$ is a very natural likelihood based measure or index on the unit interval of the extent of commonality or equality of outcome in the distribution at the modal outcome. When $LC(f) \to 0$ there is little chance of equality of outcome, when $LC(f) \to 1$ there is every chance of equality of outcome. It follows that its complement, $II(f) = 1 - LC(f) = K(1 - p_{fk^*})/(K - 1)$ is an intuitive likelihood-based measure of the extent of inequality of outcome[9]. Unfortunately, it is not responsive to variation in spread in the rest of the distribution in the sense that a marginal shift in mass from k' to k'' where k' , $k'' \neq k^*$ would leave it unaltered unless the shift rendered k'' the new modal outcome. To capture this, the concept of a modal preserving spread needs to be considered. Basically g constitutes a Modal Preserving Spread of f if [3] holds and k^*remains the modal outcome of g i.e. $p_{gk^*} = \max_k p_{gk}$.

This can be readily checked by considering $UAMBI(f,g) = \frac{\sum_{k=1}^{K}((G_k^{U,k^*} - F_k^{U,k^*}) + (G_k^{L,k^*} - F_k^{L,k^*}))}{\sum_{k=1}^{K}\left(\left|(G_k^{U,k^*} - F_k^{U,k^*})\right| + \left|(G_k^{L,k^*} - F_k^{L,k^*})\right|\right)}$, when $UAMBI(f,g)$ = 1, distribution g constitutes an unambiguous Modal Preserving Spread of distribution f. Furthermore, given dispersion from the focus point k^* is maximized when k^*/K mass is allocated to the lowest outcome and $\frac{(K-k^*)}{K}$ is allocated to the highest outcome:

[9] Indeed, in the unordered categorical world IC and II provide equally useful indices of commonality and inequality.

$$0 \leq IMPS(g,f) = \frac{\sum_{k=1}^{K}((G_k^{U,k^*} - F_k^{U,k^*}) + (G_k^{L,k^*} - F_k^{L,k^*}))}{\left(\frac{k^* \sum_{i=1}^{k^*-1} i}{K} + \frac{(K-k^*)\sum_{i=k^*+1}^{K}(i-k^*)}{K} - \sum_{k=1}^{K}\left(F_k^{U,k^*} + F_k^{L,k^*}\right)\right)} \leq 1$$

provides an index measure on the unit interval of the extent of increased Modally Focused relative spread or inequality associated with a move from f to g.

A Modally Focused Inequality Index for Ordered Categorical Data.

Suppose f^e was the distribution of a completely equal society with all agents enjoying outcome k^*, then $p_{fk^*} = 1$ and $p_{fk} = 0 \;\forall\; k \neq k^*$ so that $F_k^{U,k^*} = 0$ and $F_k^{L,k^*} = 0 \;\forall\; k$, then $IMPS(g, f^e)$ becomes:

$$IMPS(g,f^e) = \frac{\sum_{k=1}^{K}(G_k^{U,k^*} + G_k^{L,k^*})}{\left(\frac{(k^*-1)\sum_{i=1}^{k^*-1} i}{K} + \frac{(K-k^*)\sum_{i=k^*+1}^{K}(i-k^*)}{K}\right)} = MFI(g) \qquad [4]$$

[4] corresponds to a measure of the extent of inequality inherent in the ordered categorical distribution g relative to a state of complete equality at the category most likely to command unanimous membership and thus provides a measure, $MFI(g)$, of the Modally Focused Inequality inherent in distribution g. Let the k^* Focussed Probabilistic Distance vector $\underline{G}^{PD,k^*}$, recording the chance of being in the collection of categories successively further distanced from k^* , be given by:

$$\underline{G}^{PD,k^*} = \begin{bmatrix} G_1^L \\ . \\ G_{k^*-1}^L \\ 0 \\ G_{k^*+1}^U \\ . \\ G_{K-k^*}^U \end{bmatrix}$$

Note that the Probabilistic Distance function is an increasing function of the categorical distance from the i^* category which does not depend upon arbitrary attribution of value to a category in the form of a scale. Letting $\varphi(K,k^*) = \left(\frac{k^* \sum_{i=1}^{k^*-1} i}{K} + \frac{(K-k^*)\sum_{i=,k^*+1}^{K}(i-k^*)}{K}\right)$ then, given a K dimensioned unit vector d, $MFI(g)$ may be written as:

$$MFI(g,k^*) = \frac{1}{\varphi(K,k^*)} d' \underline{G}^{PD,k^*} \qquad [5]$$

Inference.

Following Rao (2009), given an independent random sample of size n, $\widehat{\underline{p}_g}$, the estimator of the vector of outcome probabilities $\underline{p}_g$ is multivariate normal:

$$\sqrt{n}\left(\widehat{\underline{p_g}} - \underline{p_g}\right) \sim N\left(\underline{0}, V_g\right)$$

where:

$$V_g = \begin{bmatrix} p_{1,g} & 0 & 0 & . & 0 \\ 0 & p_{2,g} & 0 & . & 0 \\ 0 & 0 & p_{3,g} & . & 0 \\ . & . & . & . & 0 \\ 0 & 0 & 0 & . & p_{K,g} \end{bmatrix} - \begin{bmatrix} p_{1,g} \\ p_{2,g} \\ . \\ . \\ p_{K,g} \end{bmatrix} \begin{bmatrix} p_{1,g} & p_{2,g} & . & . & p_{K,g} \end{bmatrix}$$

Given a K dimensioned square cumulation matrix C_{k^*} with typical element $c_{i,j}\ i,j = 1,..,I$ where for $i,j < k^*, c_{i,j} = 1\ when\ j \geq i\ and\ 0$ otherwise, and for $i,j > k^*, c_{i,j} = 1\ when\ j \leq i\ and\ 0$ otherwise, all other elements of the matrix are 0 [10], and a summation vector $\underline{d}$ which is a K x 1 column of ones, then $\underline{G}^{PD,k^*} = C_{k^*}\underline{p}_g$ and $\underline{\hat{G}}_g^{PD,k^*} = C_{k^*}\widehat{\underline{p_g}}$ so that:

$$\sqrt{n}\left(\underline{\hat{G}}_g^{PD,k^*} - \underline{G}_g^{PD,k^*}\right) \sim N\left(\underline{0}, C_{k^*}V_g C_{k^*}'\right)$$

So that $\widehat{MFI}(g)$, estimates of $MFI(g)$ will be such that:

$$\sqrt{n}\left(\widehat{MFI}(g) - MFI(g)\right) \sim \sqrt{n}\frac{1}{\varphi(K,k^*)}\underline{d}'\left(\underline{\hat{G}}_g^{PD,k^*} - \underline{G}_g^{PD,k^*}\right) \sim N\left(0, \frac{1}{\varphi(K,k^*)^2}\underline{d}'C_{k^*}V_g C_{k^*}'\underline{d}\right)$$

A multivariate version.

In a multidimensional ordered categorical context, one of the attractions of the probabilistic distance approach is that, unlike the corresponding cardinal environment, the unit of measure is a probability number that is common to all dimensions. For simplicity, consider the bivariate categorical case where both dimensions are ordered with $p_{f,i,j} \geq 0 : i = 1,..,I, j = 1,..,J\ \sum_{i=1}^{I}\sum_{j=1}^{J} p_{f,i,j} = 1$ with the ordering again following the dimension indexing, cumulative and counter cumulative density functions are well defined with $F_{i,j} = \sum_{k=1}^{i}\sum_{l=1}^{j} p_{f,k,l}\ \ for\ i = 1,..,I, j = 1,..,J.$

In the modal case where k^* coordinates are $\{i^*, j^*\}$ so that $\max_{i,j} p_{f,i,j} = p_{f,i^*,j^*}$:

Let $p_{f,i^*,j}^{**} = p_{f,i^*,j}\ j = 1,..,J\ and\ \ p_{f,i,j^*}^{**} = p_{f,i,j^*}\ i = 1,..,I$

[10] As an example, for I = 6 and $k^* = 3$, C_{k^*} is of the form:

$$C_{k^*} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

$$F_{i,j}^{**} = F_{i+1,j}^{**} + p_{f,i,j}\ \forall\ i < i^*\ and\ F_{i,j}^{**} = F_{i,j}^{**} + p_{f,i,j}\ \forall\ i > i^*, \forall j = 1,..,J$$

$$F_{i,j}^{\mathrm{L}k^*} = F_{i,j+1}^{\mathrm{L}k^*} + F_{i,j}^{**}\ \forall\ j < j^*\ and\ F_{i,j+1}^{\mathrm{U}k^*} = F_{i,j}^{\mathrm{U}k^*} + F_{i,j}^{**}\ \forall\ j > j^*, i = 1,..,I$$

Again, when p_{f,i^*,j^*} is viewed as the likelihood that the whole population resides in outcome $\{i^*, j^*\}$, $IC(f) = \left(IJp_{f,i^*,j^*} - 1\right)/(IJ - 1)$ is a very natural measure or index on the unit interval of the commonality or equality of outcome in the distribution, so that its complement, $II(f) = IJ\left(1 - p_{f,i^*,j^*}\right)/(IJ - 1)$ provides an intuitive likelihood based measure of inequality of outcome and it is an equally useful index of such in unordered categorical paradigms.

The corresponding 2-dimensional version of [4] is given by:

$$MFI(g) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(G_{i,j}^{U,k^*} + G_{i,j}^{L,k^*}\right)}{\left(\frac{j^*i^*\sum_{i=1}^{i^*-1}\sum_{j=1}^{j^*} ij}{IJ} + \frac{(IJ - j^*i^*)\sum_{i=i^*+1}^{I}\sum_{j=j^*+1}^{J}(ij - i^*j^*)}{IJ}\right)}$$

Appropriately vectorized versions of the I x J matrices $G_{.,.}^{U,k^*}$ and $G_{.,.}^{L,k^*}$ and their corresponding IJ square cumulation matrix $C_{i^*j^*}$ can be constructed to form the i^*, j^* Focused Probabilistic Distance vector $\underline{G}^{PD,i^*,j^*}$, recording the chance of being in the collection of categories successively further distanced from i^*, j^*. Then, given an IJ dimensioned unit vector d, $MFI(g)$ may be written as:

$$MFI(g) = \frac{1}{\varphi(IJ, i^*, j^*)} d'\underline{G}^{PD,i^*,j^*} \qquad [6]$$

Standardization.

All that remains is to standardize $MFI(g)$ with an appropriate probability-based distance measure factor to render it unit free. Analogous to the continuous paradigm formulation of the mean as the integral of the survival function over the range of x (see [1] above) the sum of the SF values over all categories $\sum_{k=1}^{K}(1 - G(k))$ could be considered so that:

$$OCCOV = MFI(g, k^*)/\left(\sum_{k=1}^{K}(1 - G(k))\right) \qquad [7]$$

would provide an ordered categorical Coefficient of Variation analogue appropriate to for the situation at hand[11].

[11] [7] is clearly dependent upon K this is of no consequence when variates with a common K are being compared, but when variates with different K's are being compared their respective values of [7] should be rescaled by $\gamma(K) = {}^{(K-1)}/_{K}$

The inverse of $OCCOV$ is the ordered categorical paradigm equivalent of the Sharpe Ratio which is a risk or uncertainty adjusted average returns measure. Thus, it can be viewed as an outcome level measure diluted by a measure of uncertainty surrounding outcome levels.

As for inference with respect to the denominator, given $G(k^*) = \sum_{i=1}^{k^*} p_{\mathrm{i},g}$, letting h be a $Kx1$ vector whose first k^* elements are ones and the rest zeros, $G(k^*) = h'\underline{p_g}$ and $\hat{G}(k^*) = h'\widehat{\underline{p_g}}$, so that $\sqrt{n}(\hat{G}(k^*) - G(k^*)) \sim N\big(0, h'V_g h\big)$ and similarly for the Survival Function Value. For $\sum_{k=1}^{K}(1 - G(k))$, construct a $K\ x\ K$ matrix S whose above diagonal elements are ones with all other elements zeros, then $\sum_{k=1}^{K}(1 - G(k)) = \underline{d}'S\underline{p_g}$ so that in a similar fashion:

$$\sqrt{n}\left(\sum_{k=1}^{K}(1 - \hat{G}(k)) - \sum_{k=1}^{K}(1 - G(k))\right) \sim N\big(0, \underline{d}'SV_g S'\big)\underline{d}$$

Axiomatics.

Suppose that $\underline{x}$ is the n x 1 dimensioned list of the category locations of n sampled individuals upon which the estimates of $\hat{p}_{fk}, k = 1,..,K$ are based, then $OCCOV$ is readily shown to satisfy the axioms of Anonymity (i.e. it is independent of the ordering of the list $\underline{x}$); Scale Invariance (it is independent of any arbitrary scale accorded the categories) and Population Independence (it will not change when the population is replicated and added to itself) it can also be shown to satisfy a weak version of the Pigou-Dalton Transfer Principal (when the presence in a higher category is reduced by 1 and the presence in a lower category increased by 1 without altering the mode, it will not increase). However, $OCCOV$ is not independent of K, the number of categories. The maximum value the numerator could take on is $0.25(K^2 - 1)$ and the maximum value the denominator could take on is $K - 1$ so the value of $OCCOV$ could be of the order of $0.25(K + 1)$. While this is of no consequence when groups are being compared over the same number of categories, it does need attention when groups are being compared over different numbers of categories so that when comparing groups based upon different numbers of categories, it may be prudent to multiply $OCCOV$ by $1/(K + 1)$ for comparison purposes.

3. An Example: Relative Variation in the distribution of Human Resources across the gender divide.

A nations Human Resource Stock (HRS), the aggregation of its constituent agents HRS, is an amorphous amalgam of their Embodied Human Capital and Cumulated Experience. The fact that males and females face different labour market and life cycle circumstances and have different knowledge acquisition traits suggests that the nations HRS has been acquired and employed differently across the gender divide (Goldin 2014) and differences between the genders of the within gender level and variability of its possession is of interest. To this end an analysis of the corresponding Coefficient of Variation can be informative regarding the relative variation and the Sharpe ratio can reveal something about the uncertainty diluted level of human resources.

Assessing the levels is difficult since both components are fundamentally latent and unobservable. Experience - the agents productivity enhancing skills acquired by practice and learning by doing - can be proxied for by the passage of time or the recorded age group of the individual. Embodied Human Capital – the agents' education and training augmented innate abilities– can be proxied for by the Education and Training level they have received. Both proxies are ordered categorical variates and, beyond the afore-mentioned issues associated with using and combining artificially attributed cardinal scales to ordinal variates, their combination in some simple algebraic form is problematic. To examine the progress of Human Resource Stocks in Canada, data on the age, education and training status of individuals have been drawn from the Census of Canada Individual Files for the years 2006 and 2016[12]. The joint probability distributions (PDF) and survival functions (SF) over experience and education and training level groups for Canadian Males and Females in 2006 and 2016 are reported in the appendix.

Table 1. Bivariate Ordered Categorical Coefficient of Variation

	Modal Experience, Training Location	Variation	Location Value	$OCCOV$	Sharpe
Females 2006	3,3	3.0662	18.2842	0.1677	5.9630
Females 2016	4,2	3.6889	19.4335	0.1865	5.3619
Males 2006	3,3	3.0592	17.5026	0.1748	5.7208
Males 2016	1,2	10.5715	18.1212	0.5834	1.7140

[12] All agents over the age of 19 who received an income and reported age and educational status were included in the study resulting in 608538 observations in 2006 (312405 of which were female) and 610346 in 2016 (326676 of which were female). An individual's experience is proxied for by their age group category with 20-29, 30-39, 40-49, 50-59, 60-69 and ≥ 70 being the designated experience categories. The education and training embodied human capital levels are based on 5 ordered categories:- EDU1: Did not finish high school, EDU2: Completed High school, EDU3: Trade or Apprentice certification or University certification or diploma below bachelor degree level and EDU5: University certificate or diploma bachelor level and above, including masters and doctorates. EDU3 and 40-49 age group were deemed the Sufficient Human Resource level.

Table 1 reports $OCCOV3$ and Sharpe Ratios together with their components for the joint density for Males and Females in 2006 and 2016. What may be gleaned from Table 1 is that relative variation of human resource stocks increased for both Females and Males over the decade, but much more so for Males, a result of the substantial shift downwards in the Male modal location engendering a substantial increase in relative variation and a sharp reduction in the uncertainty moderated level of human resources.

Table 2. Education Ordered Categorical Coefficient of Variation

		EDU1	EDU2	EDU3	EDU4	EDU5	$OCCOV$	Sharpe
Females 2006	PDF	0.2104	0.2792	**0.3109**	0.0570	0.1425	0.6346	1.5758
	Survival Function	0.7896	0.5104	0.1995	0.1425	0.0000		
Females 2016	PDF	0.1681	**0.3035**	0.2751	0.0399	0.2134	0.6367	1.5706
	Survival Function	0.8319	0.5284	0.2533	0.2134	0.0000		
Males 2006	PDF	0.2139	0.2591	**0.3502**	0.0446	0.1322	0.6139	1.6289
	Survival Function	0.7861	0.5270	0.1768	0.1321	0.0000		
Males 2016	PDF	0.2010	**0.3374**	0.2311	0.0314	0.1992	0.6462	1.5475
	Survival Function	0.7990	0.4616	0.2306	0.1992	0.0000		

Table 3. Experience Ordered Categorical Coefficient of Variation

		20-29	30-39	40-49	50-59	60-69	> 69	$OCCOV$	Sharpe
2006 Female	PDF	0.1689	0.1749	**0.2169**	0.1751	0.1203	0.1439	0.5825	1.7167
	Survival Func	0.8311	0.6562	0.4393	0.2642	0.1439	0.0000		
2016 Female	PDF	0.1673	0.1625	0.1647	**0.1921**	0.1609	0.1525	0.5891	1.6975
	Survival Func	0.8327	0.6702	0.5054	0.3133	0.1525	0.0000		
2006 Male	PDF	0.1807	0.1782	**0.2232**	0.1872	0.1195	0.1113	0.5853	1.7085
	Survival Func	0.8193	0.6412	0.4180	0.2308	0.1113	0.0000		
2016 Male	PDF	0.1904	0.1689	0.1685	**0.1959**	0.1537	0.1225	0.6360	1.5723
	Survival Func	0.8096	0.6407	0.4722	0.2763	0.1225	0.0000		

When the marginal Education and Experience distributions are considered in Tables 2 and 3 respectively, a greater increase in relative variation in both Education and Experience for males relative to females can be observed in both dimensions. Notice that, when viewed separately, the individual dimension mode changes over time are similar across genders (mode levels lowering in the case of education and increasing in the case of experience), a pattern that is not reflected in the bivariate distribution (increasing experience, decreasing education for women, both decreasing for men).

The relative variation in education as apposed to experience outcomes is examined in table 4. To make this comparison the ordered categorical coefficient of variation needs to be adjusted bye the number of categories in the respective variates. When this is done it can be observed that with the exception of males in 2016 education has greater relative variation phone does experience so the position is reversed for 2016 males.

Table 4. Education and Experience comparison.

	$OCCOV$	K adjusted $OCCOV$	K adjusted Sharpe
Education 2006 Female	0.6346	0.5077	1.9697
Education 2016 Female	0.6367	0.5094	1.9632
Education 2006 Male	0.6139	0.4911	2.0362
Education 2016 Male	0.6462	0.5170	1.9344
Experience 2006 Female	0.5825	0.4854	2.0601
Experience 2016 Female	0.5891	0.4909	2.0370
Experience 2006 Male	0.5853	0.4878	2.0502
Experience 2016 Male	0.6360	0.5300	1.8868

4. Conclusions.

By invoking the notion of probabilistic distance and developing a measure of level analogous to the mean in cardinal paradigms, it is possible to measure the relative ordinal variation in a population which is unit free and comparable across populations In spite of the lack of cardinal measure. Furthermore, the measures are easily implemented in multidimensional environments. An exemplifying application of the measure to examine the relative variability in human resources across the gender divide in 21st century Canada revealed substantial differences. Both genders increased in relative variation over the decade with males exhibiting greater variation than females with the decanal increase being much greater for males. Generally, relative variation in education and training exceeded that of experience.

References.

Aerts, S., Haesbroeck, G. and C. Ruwet (2015) “Multivariate coefficients of variation Comparison and influence functions” Journal of Multivariate Analysis. 142 183-198.

Albert, A., and L. Zhang (2010) “A novel definition of the multivariate coefficient of variation” Biometrical Journal, Volume 52, 5, 667-675

Allison, R. A. and J. E. Foster (2004). Measuring health inequality using qualitative data. Journal of Health Economics 23, 505- 24.

Anderson G.J. (2004) “Toward an Empirical Analysis of Polarization” Journal of Econometrics 122 1-26

Anderson G.J. and G. Yalonetzky (2023) “The mode and the measurement of inequality with single and multiple ordinal variables” Mimeo Economics Department University of Toronto.

Bedeian, A. G., & Mossholder, K. W. (2000). On the Use of the Coefficient of Variation as a Measure of Diversity. Organizational Research Methods, 3(3), 285–297

Blair, J. and M.G. Lacy (2000) “Statistics of Ordinal Variation” Sociological Methods and Research 28 251-280.

Bond T.N. and K. Lang (2019) "The Sad Truth about Happiness Scales," Journal of Political Economy 127, no. 4: 1629-1640.

Champernowne, D. G.; Cowell, F. A. (1999). Economic Inequality and Income Distribution. Cambridge University Press.

Galton, F. (1894), Natural Inheritance (5th ed.), New York: Macmillan and Company.

Goldin, C. (2014) “A Grand Gender Convergence: Its Last Chapter” American Economic Review 2014, 104(4): 1–30.

Zahra Jalilibal, Amirhossein Amiri, Philippe Castagliola, Michael B.C. Khoo. Monitoring the coefficient of variation: A literature review. Computers & Industrial Engineering, 2021, 161, pp.107600. ff10.1016/j.cie.2021.107600ff. ffhal-03336179f

Kobus, M. (2015). Polarization measurement for ordinal data. Journal of Economic Inequality 13, 275-297.

Kobus M. and Kurek R. (2019) “Multidimensiona1 polarization for ordinal data” The Journal of Economic Inequality 17: 301-317.

Kvalseth, T.O. (2017) “Coefficient of Variation: the Second-Order Alternative” Journal of Applied Statistics 44 402-415.

Mendelson, H. (1987). Quantile- preserving spread. Journal of Economic Theory 42, 334-51.

Pearson K. (1896) “Mathematical contributions to the theory of evolution: Regression, heredity and panmixia” Philos. Transactions of the Royal Society London. A, Math. Phys. Eng. Sci 187, 253– 318.

Reyment, R.A. (1960) Studies on Nigerian upper Cretaceous and lower tertiary ostracoda: part 1 Senonian and maastrichtian ostracoda. Stockolm Contributions in Geology 7 1-238.

Schroder, C. and Yitzhaki, S. (2017). Revisiting the evidence for cardinal treatment of ordinal variables. European Economic Review, 92:337-358.

Sharpe, W.F. (1964) “Capital Asset Prices: A theory of Market Equilibrium Under Conditions of Risk” Journal of Finance 428-442.

Sharpe, W. (1994) “The Sharpe ratio”, Journal of Portfolio Management, Vol. 21, No. 1, Fall, pp. 49-58.

Sortino F.A. and L.N. Price (1994) “Performance Measurement in a Downside Risk Framework” The Journal of Investing 3 59-64.

Van Valen L. (1974) “multivariate structural statistics in Natural History” Journal of Theoretical Biology 45 235-247.

Voinov V.G. and M.S. Nikulin (1966) “Unbiased Estimators and their Applications” Multivariate Case Vol. 2 Kluwer.

Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting Risk Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation. *Psychological Review, 111*(2), 430–445.

Apppendix. Joint PDF's and Survival Functions.

		Joint PDF					Survival Function				
		EDU1	EDU2	EDU3	EDU4	EDU5	EDU1	EDU2	EDU3	EDU4	EDU5
Females 2006	20-29	0.0181	0.0552	0.0543	0.0077	0.0337	0.9819	0.9268	0.8725	0.8648	0.8311
	30-39	0.0182	0.0390	0.0666	0.0105	0.0405	0.9637	0.8696	0.7487	0.7305	0.6562
	40-49	0.0294	0.0628	0.0795	0.0129	0.0323	0.9343	0.7773	0.5769	0.5459	0.4393
	50-59	0.0320	0.0540	0.0557	0.0114	0.0219	0.9023	0.6913	0.4351	0.3928	0.2642
	60-69	0.0408	0.0321	0.0304	0.0081	0.0089	0.8615	0.6184	0.3318	0.2814	0.1439
	> 69	0.0719	0.0361	0.0243	0.0066	0.0050	0.7896	0.5104	0.1995	0.1425	0.0000
Females 2016	20-29	0.0137	0.0588	0.0436	0.0052	0.0461	0.9863	0.9275	0.8839	0.8787	0.8327
	30-39	0.0135	0.0364	0.0511	0.0067	0.0548	0.9728	0.8775	0.7828	0.7710	0.6702
	40-49	0.0166	0.0404	0.0544	0.0075	0.0459	0.9562	0.8206	0.6714	0.6522	0.5054
	50-59	0.0272	0.0638	0.0597	0.0080	0.0335	0.9290	0.7296	0.5209	0.4936	0.3133
	60-69	0.0332	0.0580	0.0408	0.0066	0.0223	0.8958	0.6385	0.3890	0.3551	0.1525
	> 69	0.0639	0.0461	0.0256	0.0060	0.0108	0.8319	0.5284	0.2533	0.2134	0.0000
Males 2006	20-29	0.0265	0.0673	0.0565	0.0067	0.0238	0.9736	0.9063	0.8498	0.8431	0.8193
	30-39	0.0244	0.0444	0.0675	0.0083	0.0335	0.9491	0.8374	0.7135	0.6985	0.6412
	40-49	0.0390	0.0557	0.0873	0.0100	0.0311	0.9101	0.7427	0.5313	0.5063	0.4180
	50-59	0.0372	0.0473	0.0683	0.0093	0.0251	0.8728	0.6582	0.3785	0.3442	0.2308
	60-69	0.0374	0.0245	0.0400	0.0059	0.0117	0.8354	0.5962	0.2766	0.2364	0.1113
	> 69	0.0493	0.0199	0.0306	0.0044	0.0071	0.7861	0.5270	0.1768	0.1321	0.0000
Males 2016	20-29	0.0236	0.0855	0.0397	0.0041	0.0375	0.9764	0.8909	0.8512	0.8471	0.8096
	30-39	0.0223	0.0525	0.0454	0.0051	0.0436	0.9540	0.8160	0.7309	0.7217	0.6407
	40-49	0.0246	0.0524	0.0446	0.0061	0.0408	0.9295	0.7390	0.6093	0.5940	0.4722
	50-59	0.0412	0.0634	0.0494	0.0062	0.0356	0.8882	0.6344	0.4552	0.4337	0.2763
	60-69	0.0373	0.0507	0.0333	0.0055	0.0269	0.8509	0.5464	0.3339	0.3068	0.1225
	> 69	0.0520	0.0328	0.0186	0.0043	0.0149	0.7990	0.4616	0.2306	0.1992	0.0000