Qual Quant (2014) 48:2647–2656
DOI 10.1007/s11135-013-9914-5

The co-occurrence of self-observed norm-conforming behavior, reduction of zero observations and remaining measurement quality

Johannes A. Landsheer

Received: 9 February 2013 / Accepted: 25 July 2013 / Published online: 10 August 2013

Abstract Norm-violating behavior is characterized by clear social norms which prescribe the non-occurrence of that behavior. From the theoretical framework of Allport it is derived that specifically norm-conformation is consistent, while violating norms is expected to be inconsistent and more circumstantial. This is in contrast to test-theoretic approaches of delinquent behavior that assume that various norm-violating responses form a consistent answer pattern that is scalable and reliable. In this study we study the inter-correlations, scalability and reliability of norm-violating responses and their relation with the reduction of zero observations. In concordance with Allport's view it is expected that different norm-violating self-report items have limited interrelatedness and are limited in scalability and reliability in the norm-violating sub-population. The NLSY98 self-report data show that a large majority of respondents (69 %) conform systematically to all ten different norms, while only nine percent admits more than two different violations. The results show that in subsamples of norm-violating respondents, the correlations between items become closer to zero, dependent on the amount of zero reduction. Furthermore, both Loevinger's H coefficient of scalability and scale reliability become unsatisfactorily low, when 35 % or more strict norm-conforming subjects are removed.

Keywords Norm violations · Zero inflation · Measurement quality · Criminology · Scalability · Reliability

1 Introduction

When a collection of norm-violating behaviors are studied, most of these behaviors are positively correlated. This forms the basis for the widespread application of test theory in the field of criminology, either classical test theory (CTT) or item response theory (IRT). Applications of test theory concerning norm-violating behavior, such as criminal behavior

J. A. Landsheer (✉)
Department of Methodology and Statistics, Faculty of Social Sciences, Utrecht University,
Padualaan 14, 3584 CH Utrecht, The Netherlands
e-mail: j.a.landsheer@uu.nl

can be found in various publications (Huizinga and Elliott 1986; Osgood et al. 2002; Piquero et al. 2002; Raudenbush et al. 2003). A self-report inventory consists of a list of questions whose answers are supposed to reflect the behavior of the respondents (Goldstein and Wood 1989), in this case norm-violating behavior. The answers presuppose a probabilistic relationship with an underlying characteristic or trait of the respondent. The answers would then provide an opportunity to classify persons in the degree to which the person has the underlying property or trait. Given the positive correlations between self-reported norm violations, it can be assumed that norm violations co-occur. When somebody steals, other concurrent norm violations of the same person, such as destroying property or selling drugs, would allow describing a more general pattern, offering a possibility to predict one behavior on the basis of another. This assumption of co-occurrence forms the foundation of the self-report measurement of delinquency and these scales are based on a variety of different delinquent behaviors.

Typically, the resulting variables have zero as the only norm-conforming score. These variables have characteristically a large number of norm-conforming observations of zero and a limited number of norm-violating scores larger than zero and are therefore skewed. This kind of variables has specific statistical properties: they have a mode at zero, a median at zero (that is, more than 50 % zero scores), a relative high variance or dispersion, and a strongly skewed distribution. Furthermore, these variables are most often positively correlated (Osgood et al. 2002). The correlation of parallel observations of norm-violating behavior is especially relevant for constructing a test, when several of these variables are used to construct scales.

Recently, several statistical models have been developed that can be used for the statistical analysis of this kind of data. The statistical view on such high numbers of observed zeros is "zero inflation" (Lambert 1992; Cameron and Trivedi 1998). This general view is based on the assumption of two subpopulations, one that is "at risk" or has to deal with "uncommon circumstances", which produces observations in the positive range and a second subpopulation, which produces solely zero observations. Members of this subpopulation are "never at risk" or have to deal only with "common circumstances" and offer solely zero counts as a consequence (Lambert 1992). Nagin and Land (1993) have extended this to a finite-mixture trajectory model for delinquent and criminal behavior, assuming mixed distributions of groups of respondents or latent classes that develop their behavior according to distinguishable delinquent trajectories. When applied to a single cross-section, the Nagin and Land model reduces to a latent class model, using zero-inflated Poisson or negative Binomial distributions in which a large part of the zero scores form a separate class.

Although the zero-inflation approach and other models using mixed distributions have been widely applied over the past decades in criminological research, several statisticians have expressed their concerns. A general consideration is that this model does not always fit the data very well (Slymen et al. 2006). Min and Agresti (2005) have formulated a variety of concerns: the zero-inflation model may be unsuitable to fit zero-inflated data; it may be unreliable when fitting zero-inflated count data, and the random effect model requires that the data is zero-inflated at every level of the covariates, which may be unrealistic. Bauer and Curran (2003) warned that the identification of classes may be influenced by skewed distributed data. More recently, Ghosh et al. (2012) found that zero-inflation models struggle to explain extreme incidence of zeros, and may even struggle when the proportion is not extreme.

Perumean-Chaney et al. (2012) address the problem that zero counts may originate from various mechanisms: Firstly, the zeros may be "true" values reflecting an absence of the event of interest. But secondly, zero counts may reflect behavior of individuals who are not

at risk because a suitable possibility does not exist. And thirdly, zero counts may be the result of under-reporting. In addition to these mechanisms we add behavior-specific and circumstances specific zero counts: a fourth mechanism is that some individuals may simply choose to violate norms in one specific respect, while conforming to other norms. A fifth mechanism is that specific circumstances may allow for certain norm-violations only. These latter two mechanisms explain that norm-violators have a non-zero score on some behavior and a zero score on others, making norm violations less consistent.

Allport's J-curve hypothesis of conforming behavior offers an interesting definition of the specific properties of this behavior (Allport 1934, 1939). Allport noticed that this kind of observations often have the shape of a J lying down, hence the name J-curved variables. He sees the J-curves of conforming behavior as a summation of four component distributions, produced, respectively, by conformity-making agencies, common biological tendencies, personality differences, and chance. The existence of clear social norm (laws) and strong social agencies (police; judicial apparatus) assure that a person conforms to the desired mode and enforces norm-conforming behavior. Allport's theory allows us to formulate various expectations concerning the correspondence of conforming behavior. Strong conformity producing agencies (laws, police force/judiciary apparatus) are external and equally active for different norms and behaviors.

Both the J-curve theory and zero-inflation theory assume a mixed distribution. There are however differences: Allport assumes a strong representation of the norm-conforming score and sees the large number of norm-conforming scores as inherent of norms and the strong social agencies that enforce norm-conformation. In zero-inflation theory, the assumption of an "at risk" subpopulation points in the direction of circumstances and perhaps personality characteristics which may have differential influences in behavioral expression and that allow for the prediction of norm violations. Both the J-curve hypothesis and the zero-inflation model assume a subpopulation (respectively norm-conforming and "not at risk"), distinct from the other subpopulation (respectively norm-violating and "at risk"). Correlations and/or regression coefficients found in the data collection without distinction of the two subpopulations may be non-existent or have other values in the two separated subpopulations.

The basic question in this study is whether the behavioral pattern in the group of norm violators is sufficiently consistent and their responses are sufficiently correlated. Furthermore, we expect that self-report responses of norm-violations are insufficiently scalable and reliable. In this paper, self-reported norm violations of adolescents are used to demonstrate the empirical properties of these responses and most specifically the correlations between these variables. We expect that the earlier mentioned correlations between norm violations that are commonly found in an undivided sample of the general population cannot be found in the subsample of norm violators.

Our first hypothesis is that the correlations of any two different norm-violating behaviors are strongly dependent on the zero scores. Our second hypothesis is that norm violations are limited in their consistency across norms and behavior and are therefore insufficiently correlated within the sample that is considered as the norm violating respondents. Our last hypothesis is that the responses of norm-violations are insufficiently scalable and reliable in the subpopulation that remains when the norm-conforming subpopulation has been removed partially.

2 Methods

In this paper we study several aspects of measurements of norm-violating behavior. The raw correspondence found in different norm-violating behaviors is reported. To enhance the

understanding of this kind of variables the empirical properties are examined when zero scores are removed systematically. We show the effects of reduction of the systematic zero sum score on the interrelatedness of self-reported norm violations. Furthermore, we show the effects on scalability and reliability when systematic zeros are removed. As an index of scalability, Loevinger's H is used, which indicates the scalability of the total scale. A H coefficient <0.30 indicates that the indicators do not form a scale. When the H value is around 0.30 the indicators are considered to form a weak scale (Loevinger 1948; van der Ark et al. 2008). As an index of scale reliability we have used Cronbach's Alpha. Although there is substantial discussion about the interpretation of this coefficient (Cortina 1993; Sijtsma 2009), a low Alpha indicates item-specific variance or item uniqueness and low communalities (Cortina 1993) and therefore limited reliability.

3 Data

The data set is a part of the National Longitudinal Survey of Youth study, executed in the USA, which has started in 1997. The NLSY is a national longitudinal survey which focuses on labor market experiences of American men and women (Moore et al. 2000). This dataset is chosen for its various qualities, among which the size and the relatively low number of missing values. The data can be acquired from the web site of the U.S. Department of Labor, Bureau of Labor Statistics (http://www.bls.gov/nls/home.htm).

The NLSY in 1998 studied 8,984 adolescents aged 12–16, who were just beginning the school-to-work transition. Data are collected on a large number of subjects, including delinquent activities. We have used the data of ten self-report delinquency items (see Table 1), collected in 1998. Each item represents a specific norm violation.

The data has been prepared by recoding all frequencies larger than zero to one. This forces a maximal correspondence for the frequencies of norm violations, removes the differences in high scores and hence the distortions as a consequence of outliers. Dichotomization is often used for the self-report measurement of delinquency (Hindelang et al. 1981; Piquero et al. 2002; Raudenbush et al. 2003). Furthermore, we have removed all missing values by list wise deletion. This resulted in a reduction of 656 incomplete responses to a total of

Table 1 Delinquency items of NLSY97

No.	Question	% > 0	1998			
			Raw		Dichotomous	
			$\bar{X}$	σ^2	$\bar{X}$	σ^2
1	Used marijuana during school or work	4.47	.38	7.27	.04	.04
2	Used cocaine or other drugs	5.15	5.35	1813.11	.05	.05
3	Used cocaine or other drugs during school or work	1.51	.16	4.52	.02	.01
4	Ran away	6.38	.16	1.25	.06	.06
5	Destroyed property	10.90	.66	25.02	.11	.10
6	Stolen < $50	11.92	.97	38.81	.12	.11
7	Stolen > $50	3.84	.38	19.50	.04	.04
8	Committed other property crimes	3.27	.28	1.72	.03	.03
9	Attacked or assaulted someone	11.77	.54	14.01	.12	.10
10	Sold drugs	5.88	1.30	8.24	.06	.06

Basic format: "How many times have you..?"

Table 2 Raw correlations after dichotomization

Item	1	2	3	4	5	6	7	8	9	10
1	1.00	.27	.39	.16	.17	.20	.20	.18	.17	.38
2	.27	1.00	.52	.17	.21	.22	.21	.21	.17	.40
3	.39	.52	1.00	.14	.14	.16	.21	.15	.13	.28
4	.16	.17	.14	1.00	.16	.16	.14	.11	.20	.18
5	.17	.21	.14	.16	1.00	.35	.26	.31	.28	.27
6	.20	.22	.16	.16	.35	1.00	.34	.26	.19	.29
7	.20	.21	.21	.14	.26	.34	1.00	.31	.19	.28
8	.18	.21	.15	.11	.31	.26	.31	1.00	.20	.29
9	.17	.17	.13	.20	.28	.19	.19	.20	1.00	.25
10	.38	.40	.28	.18	.27	.29	.28	.29	.25	1.00

8328. When considering the complete sample, the ten self-report items seem to be sufficient scalable (Loevinger's H is .33) and reliable (Cronbach's Alpha is .73).

4 Results

Several properties of these variables are presented in Table 1: percentages of scores larger than zero, means and variance. The raw variables have extreme high scores and are heavily skewed. Means are low and variances are relatively large. Variable 2 (drug use) is the most extreme in this respect, as a consequence of the more frequent occurrence of high scores. As described by Allport, all variables show a large amount of norm-conforming zero scores, while the variation in norm-violating scores is relatively high. Both mode and median of all variables is zero. Sixty-nine percent of all respondents provide solely zero answers on all ten items (zero sum score), while only 9 % admit more than two norm violations.

We describe here in some detail the raw correspondence in scores between the sets of variables. The Pearson product moment correlations are shown in Table 2. The correlations deviate considerably. The minimum correlation is .04, the maximum is .57 and the mean correlation is .25.

The effect of removing respondents with strict norm compliance on the correlation coefficient is considerable. Figure 1 shows the effect of reducing the number of systematic zeros (zeros on both variables) on the correlation of the first two items. Zero is reached when the systematic zeros are reduced with 84.42 %. We see a gradual diminishment of the inter-item correlation with a minimum when the absolute correlation becomes zero. This relationship can be found by all item combinations and is in fact completely systematic. In Appendix this systematic relationship is defined.

In Table 3 the minimization of all item combinations is presented and we can see that all item correlations are diminished by reducing the number of systematic zeros. The upper triangle shows the percentages with which systematic zeros have to be removed to reach an inter-item correlation of (near) zero. All correlations can be reduced to zero, except item 2 and 3. Both item 2 and 3 concern the use of cocaine, that is, identical behavior. This means that when we remove the group of norm-conforming respondents for any of two different norm violations, the correlations between the responses of the norm violating group is effectively reduced to zero.

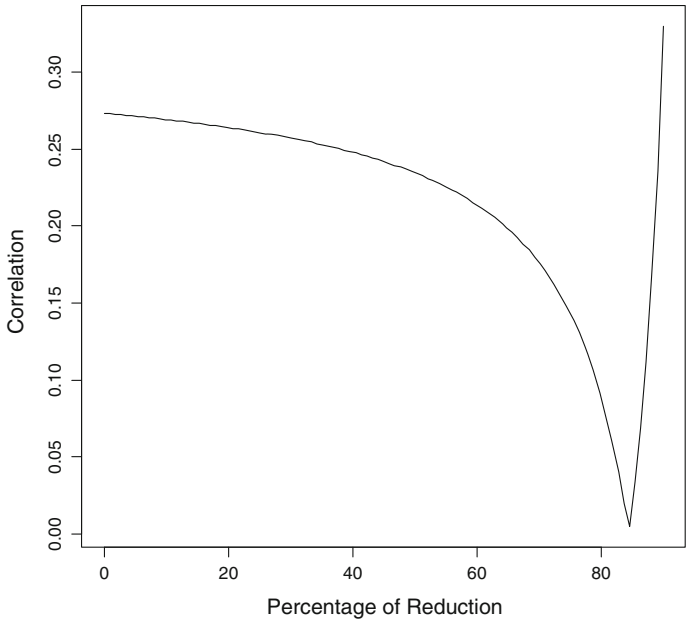

Fig. 1 The reduction of the absolute inter-item correlation of the first two self-report items

Table 3 The minimum correlation of the item combinations (lower triangle) when removing zero scores.

Item	1	2	3	4	5	6	7	8	9	10
1		84.42	93.61	74.23	69.52	71.57	82.58	82.15	68.75	87.65
2	.00		98.23	73.20	71.83	72.21	81.70	83.13	66.58	87.26
3	.00	.04		81.33	76.73	77.92	89.23	87.15	74.44	9.13
4	.00	.00	.00		63.83	62.54	72.81	69.59	67.96	73.76
5	.00	.00	.00	.00		73.01	79.00	82.96	68.67	75.85
6	.00	.00	.00	.00	.00		82.43	79.49	59.00	75.86
7	.00	.00	.00	.00	.00	.00		89.45	72.12	84.95
8	.00	.00	.00	.00	.00	.00	.00		74.81	86.46
9	.00	.00	.00	.00	.00	.00	.00	.00		73.06
10	.00	.00	.00	.00	.00	.00	.00	.00	.00	

The upper triangle shows the accompanying percentage of reduction of zero observations

Figures 2 and 3 show the effects of zero reduction on the scalability coefficient Loevinger's H and on the reliability coefficient Cronbach's alpha. It shows the results for the group that remains after a percentage of strictly norm-conforming respondents have been removed. Respondents who have a zero score on all ten items are strictly norm-conforming. In the undivided group, before any reduction of zero values, Loevinger's H is .33. This value is slightly above the minimum value of .3 for considering the items as scalable. Reduction of the zero sum observations of all ten items, result soon in a value lower that .3 (percentage of reduction 27 % or more). Further reduction of zero observations results in even lower values.

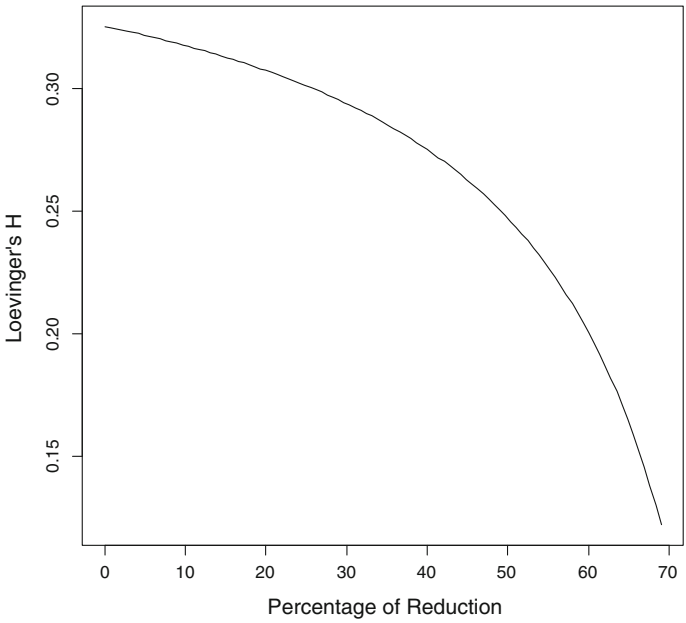

Fig. 2 The scalability of ten self-reported norm violations

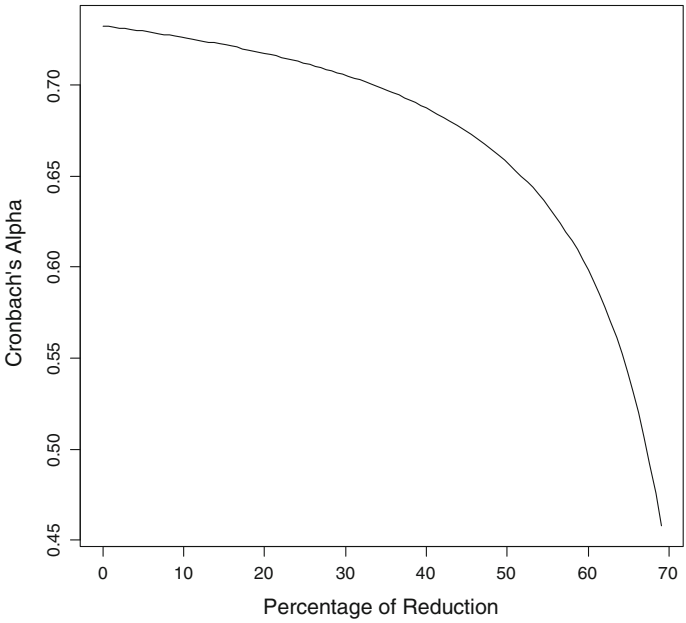

Fig. 3 The reliability of ten self-reported norm-violations after reduction of zero observations

Figure 3 shows the effects of zero reduction on the reliability coefficient Cronbach's α and is very similar to Fig. 2. In the undivided group α is .73, slightly above the minimum value (.7) for an acceptable α. Reduction leads to a lower value for α: a reduction of 35 % of

more results in a α that is $< .7$ and the scale reliability is considered as questionable. When a group of more than 60 % strictly norm-conforming respondents has been split off, the scale in the remaining group has an α lower than .6 and the scale reliability is considered as poor.

5 Conclusions

There is a large subsample (69 % of the total sample) that does not report any violation of ten different norms. We have shown that the correlation of any two different norm violating behaviors is strongly dependent on the estimation of the amount of zero scores that are considered as "true" zero scores. The correlations of any two norm violations is reduced when removing zero scores and even removing a moderated amount of zero scores (35 %) results in a substantial reduction of these inter correlations. Lastly, we have shown that the self-report items become insufficiently scalable and reliable within the group of norm violators, even when the group of strict norm conforming respondents is estimated to be no larger than 35 % of the total sample.

6 Discussion

In the case of self-reported delinquent behavior, it makes sense to interpret the distribution of these J-curved responses as a mix of at least two different subpopulations: a large subpopulation of norm-conforming respondents and a smaller subpopulation of norm-violating respondents. As in Simpson's paradox (Simpson 1951), the distinction between the two subpopulations is essential for understanding the association of two J-curved variables. However, and in contrast to the Simpson paradox, the split is based on the counts themselves, and is concentrated on a single norm-conforming score of zero.

The results show that self-reported norm violations are limited in their consistency across norms and behaviors. This makes the assumption precarious that self-reported norm violations are predictive of other norm violations. Different norm violations may for instance occur independently, despite considerable correlations found in the undivided sample. Including or excluding a large subpopulation of norm-conforming respondents may result in serious changes in the estimates of covariance and correlation and misinterpretation is easy. When removing a part of the group that has strict norm-conforming responses, inter-correlations between different items decrease. The inter-correlations are therefore dependent on the estimate of the size of the group that is considered as strict norm-conforming. Another part of strict zero responses may be considered as part of the norm-violating subpopulation as the consequence of either under-reporting or lack of opportunities. The correct size of the norm conforming group is difficult to estimate, as various assumptions are possible and the validity of these assumptions is difficult to prove.

It does make sense to interpret the distribution of the responses on self-report delinquency items as a mix of a subpopulation of norm-conforming respondents and a subpopulation of norm-violating respondents. A technical correction for the large number of norm-conforming responses may provide better estimates for the true correlations in the norm-violating subpopulation, but these estimates are necessarily lower. It is an interesting question for further research whether the reliability of delinquency measurement instruments may be overestimated, even though these reliabilities are in practice already low (Huizinga and Elliott 1986). If so, this is a possible explanation for the poor validity results of these scales, discussed in the reviews of Huizinga and Elliott (1986); Junger-Tas and Marshall (1999) and Thornberry and Krohn (2003).

A more substantive solution seems preferable. Allport (1939) has pointed to the desirability of measuring J-curved variables across a continuum and proposed a mirrored J-curve for the norm-conforming part of the distribution, the double J-curved model. Such an extended measurement would allow for the attribution of various scores to the norm-conforming respondents. This is often easier said than done and many variables concerning norm violations used in current research are still truncated at zero, skewed and do not continue in the direction of norm-conformation. However, Clarke (1996) successfully observed norm-conforming behavior as well as norm-violating behavior and he used a single variable for the amount of adherence to a speed limit. In the case of self-reported delinquent behavior, it is less clear how the conforming part of the distribution should be defined and observed so that differentiation of the norm-conforming scores is obtained. Nevertheless, it is desirable to attribute non-delinquents other scores than zero that would reflect the distance they have from respondent who violate norms. Future research may clarify whether current covariates of self-reported delinquency truly differentiate within the group of norm-violators.

Appendix

The relationship between mean, variance, covariance and correlation with and without a proportion zero-inflation of π.

Let there be a zero-inflated variable with n observations, a proportion zero-inflation of π and an underlying not inflated distribution with $\tilde{n}$ observations, mean $\tilde{\mu}$ and variance $\tilde{\sigma}^2$.

$$\mu = \frac{1}{n}\sum y = \frac{1}{n}\sum \tilde{y} = \frac{\tilde{n}}{n}\tilde{\mu} = (1-\pi)\tilde{\mu}$$

Derivation of formula (3)

The variance of a zero-inflated variable y with n observations and mean μ is:

$$\sigma^2 = \frac{1}{n}\sum y^2 - \mu^2$$

$$\sigma^2 = \frac{\tilde{n}}{n}\left(\frac{1}{\tilde{n}}\sum \tilde{y}^2 - \frac{\tilde{n}}{n}\tilde{\mu}^2\right)$$

$$\sigma^2 = \frac{\tilde{n}}{n}\left(\frac{1}{\tilde{n}}\sum \tilde{y}^2 - \tilde{\mu}^2 + \frac{n}{n}\tilde{\mu}^2 - \frac{\tilde{n}}{n}\tilde{\mu}^2\right)$$

$$\sigma^2 = (1-\pi)\left(\tilde{\sigma}^2 + \frac{n-\tilde{n}}{n}\tilde{\mu}^2\right) = (1-\pi)(\tilde{\sigma}^2 + \pi\tilde{\mu}^2)$$

and

$$\tilde{\sigma}^2 = \frac{\sigma^2}{(1-\pi)} - \frac{\pi}{(1-\pi)^2}\tilde{\mu}^2$$

The covariance formula can be derived in similar fashion.

$$\sigma_{xy} = (1-\pi)\tilde{\sigma}_{xy} + \pi(1-\pi)\tilde{\mu}_x\tilde{\mu}_y.$$

and

$$\tilde{\sigma}_{xy} = \frac{\sigma_{xy}}{(1-\pi)} - \frac{\pi}{(1-\pi)^2}\tilde{\mu}_x\tilde{\mu}_y.$$

The Pearson product moment correlation is

$$\tilde{\rho}_{xy} = \frac{\tilde{\sigma}_{xy}}{\tilde{\sigma}_x\tilde{\sigma}_y}.$$

References

Allport, F.H.: The J-curve hypothesis of conforming behavior. J. Soc. Psychol. **5**, 141–183 (1934)

Allport, F.H.: Rule and custom as individual variations of behavior distributed upon a continuum of conformity. Am. J. Sociol. **44**(6), 897–921 (1939)

Bauer, D.J., Curran, P.J.: Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. Psychol. Methods **8**(3), 338 (2003)

Cameron, A.C., Trivedi, P.K.: Regression Analysis of Count Data. Cambridge University Press, Cambridge (1998)

Clarke, R.V.: The distribution of deviance and exceeding the speed limit. Br. J. Criminol. **36**(2), 169 (1996)

Cortina, J.M.: What is coefficient alpha? An examination of theory and applications. J. Appl. Psychol. **78**(1), 98 (1993)

Ghosh, S., Gelfand, A.E., Zhu, K., Clark, J.S.: The k-ZIG: Flexible modeling for zero-inflated counts. Biometrics **68**, 878–885 (2012)

Goldstein, H., Wood, R.: Five decades of item response modelling. Br. J. Math. Stat. Psychol. **42**(2), 139–167 (1989)

Hindelang, M.J., Hirschi, T., Weis, J.G.: Measuring Delinquency. Sage Publications, Beverly Hills (1981)

Huizinga, D., Elliott, D.S.: Reassessing the reliability and validity of self-report delinquency measures. J. Quant. Criminol. **2**(4), 293–327 (1986). doi:10.1007/BF01064258

Junger-Tas, J., Marshall, I.H.: The self-report methodology in crime research. Crime Justice 291–367 (1999)

Lambert, D.: Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics **34**(1), 1–14 (1992)

Loevinger, J.: The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. Psychol. Bull. **45**(6), 507 (1948)

Min, Y., Agresti, A.: Random effect models for repeated measures of zero-inflated count data. Stat. Model. **5**(1), 1 (2005)

Moore, W., Pedlow, S., Krishnamurty, P., Wolter, K., Chicago, I.L.: National Longitudinal Survey of Youth 1997 (NLSY97). National Opinion Research Center, Chicago, IL (2000)

Nagin, D.S., Land, K.C.: Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson Model. Criminology **31**(3), 327–362 (1993)

Osgood, D.W., McMorris, B.J., Potenza, M.T.: Analyzing multiple-item measures of crime and deviance I: Item response theory scaling. J. Quant. Criminol. **18**(3), 267–296 (2002)

Perumean-Chaney, S.E., Morgan, C., McDowall, D., Aban, I.: Zero-inflated and overdispersed: What's one to do? J. Stat. Comput. Simul. http://www.tandfonline.com/doi/abs/10.1080/00949655.2012.668550 (2012). Accessed 10 Jan 2013

Piquero, A.R., Macintosh, R., Hickman, M.: The validity of a self-reported delinquency scale: Comparisons across gender, age, race, and place of residence. Sociol. Methods Res. **30**(4), 492 (2002)

Raudenbush, S.W., Johnson, C., Sampson, R.J.: A multivariate, multilevel Rasch model with application to self-reported criminal behavior. Sociol. Methodol. 169–211 (2003)

Sijtsma, K.: On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika **74**(1), 107–120 (2009)

Simpson, E.H.: The interpretation of interaction in contingency tables. J. R. Stat. Soc. Ser. B Methodol. **13**(2), 238–241 (1951)

Slymen, D.J., Ayala, G.X., Arredondo, E.M., Elder, J.P.: A demonstration of modeling count data with an application to physical activity. Epidemiol. Perspect. Innov. **3**(1), 3 (2006)

Thornberry, T.P., Krohn, M.D.: Comparison of self-report and official data for measuring crime. Measurement Problems in Criminal Justice Research, pp. 43–94. National Academies Press, Washington, DC (2003)

van der Ark, L.A., Croon, M.A., Sijtsma, K.: Mokken scale analysis for dichotomous items using marginal models. Psychometrika **73**(2), 183–208 (2008)