Thomas and Hettmansperger *Journal of Statistical Distributions and Applications* (2017) 4:27
DOI 10.1186/s40488-017-0071-6

Journal of Statistical Distributions and Applications

RESEARCH Open Access

Risk ratios and Scanlan's HRX

Hoben Thomas[1]* and Thomas P. Hettmansperger[2]

*Correspondence: hxt@psu.edu
[1]Pennsylvania State University, Department of Psychology, 121 Moore Bldg, University Park, PA 16802, USA
Full list of author information is available at the end of the article

Abstract
Risk ratios are distribution function tail ratios and are widely used in health disparities research. Let A and D denote advantaged and disadvantaged populations with cdfs $F_A(x)$ and $F_D(x)$ respectively, $F_A(x) \leq F_D(x)$. Consider a selection setting where those selected have $x > c$ a critical value. Scanlan observed in empirical data that as c is lowered the failure ratio $FR(c) = F_D(c)/F_A(c)$ and success ratio $SR(c) =[1 - F_D(c)] / [1 - F_A(c)]$ can both be increasing with decreasing c, a surprising result Scanlan calls Heuristic Rule X (HRX). The consequences of HRX for disparities research have not been well understood, and appear to be often ignored. The analytical conditions for HRX to hold have not, heretofore, been proven. In the normal case, HRX holds if the variances are equal. In general, HRX is not a robust condition. Settings and distributional conditions necessary for HRX to hold are discussed, including mixture settings of normal and other variables.

Keywords: HRX, Mixture models, Risk ratios, Tail ratios, Disparity

Mathematics Subject Classification: 60E99

1 Introduction

James P. Scanlan, a Washington D.C. lawyer, has written extensively on statistical issues of discrimination for more than two decades. He has contributed prodigiously to commentaries, workshop presentations, legal briefs, news articles, web posts, lengthy letters to organizations, including the American Statistical Association, various slide files as well as several journal publications. Typically his focus is on various measures of different population disparities. Most of this work is freely available on his burgeoning website www.jpscanlan.com. Many of Scanlan's contributions have focused on health disparities and commonly associated statistical measures of disparities. His most important contributions have addressed what epidemiologists refer to as risk ratios, or relative risk.

Risk ratios are probability distribution function tail ratios, either empirical or theoretical, or sometimes functions thereof. Scanlan views risk ratios as having been improperly understood and wrongly interpreted; as he sees it, his goal is to provide a more enlightened perspective on their use. This is a core theme in his contributions. There is a very large and often controversial literature on the use of risk ratios stretching back well more than 50 years (Sheps 1958) and which is continuing (Talih and Huang 2016; Weissman et al. 2011).

The focus here is on a surprising property of risk ratios apparently discovered by and called by Scanlan Heuristic Rule X (HRX) or sometimes "Scanlan's Rule." In more recent contributions he has tended to avoid the term and focused on certain properties of HRX instead. Let F denote a distribution function lower tail, $1 - F$ the upper tail. Consider group, D, disadvantaged with respect to another group A, with group A right-shifted, relative to D with equal scales. Define the fail ratio FR and the success ratio SR with respect to variable X.

$$FR(x) = F_D(x)/F_A(x)$$

and

$$SR(x) = [1 - F_D(x)] \,/\, [1 - F_A(x)]$$

where the subscripts A and D refer to the advantaged and disadvantaged groups, respectively. Typically $SR(x) \leq 1$ while $FR(x) \geq 1$. The terms "disparity" or "disparities" refer to the relative departure of one or both of these ratios from one, where one can be regarded as "equity." The magnitudes of these disparities can trigger public policy concerns, especially claims of racism, discrimination, or bias, and consequently sometimes legal actions.

The conundrum is the following: Suppose c is a test score above which there is success, and below which there is failure. If c is shifted downward (lowered) then $SR(x)$ increases toward one. However under certain mild conditions the failure ratio, $FR(x)$ unexpectedly increases away from one. This is the property HRX.

Scanlan has argued in numerous places with forcefulness and often at great length about properties of these ratios he has long identified; however, he uses neither the above notion nor does he characterize matters in the terms we use here. Consider the graphs of pairs $(x, FR(x))$, and $(x, SR(x))$ with the risk ratio graphed against x. Scanlan would argue that the corresponding trajectories of these graphs are grossly misunderstood…even by professionals whose business, presumably, is to understand them. For example he *starts* his "Divining Difference" 1994 *Chance* article with the statement that "There are few statistical phenomena that are at once so fundamental and so widely misunderstood as the seemingly paradoxical relationship between disparities (i.e., SR and FR above)…. The misunderstanding of that relationship is responsible for immense confusion in the appraisal of a wide range of phenomena disparately affecting different demographic groups (Scanlan 1994, p. 38)".

He prefers to characterize the difficulty this way: "The main problem with existing research lies in the failure to recognize the following statistical tendency, which we'll call heuristic rule X (HRX): When two groups differ in their susceptibility to an outcome, the rarer the outcome, the greater the disparity in experiencing the outcome and the smaller the disparity in avoiding the outcome. (Scanlan 2006, p. 47)".

Readers (ourselves included) often find such statements hard to parse partly because he typically provides an insufficient framework at the outset, visual or conceptual, within which to parse the prose. Scanlan himself has come, recently, to recognize this difficulty although he appears reluctant to alter his presentation style (Scanlan 2016a, footnote 3, p. 7).

The goal here is to explore analytically and empirically HRX, and to show under what conditions HRX holds. It seems remarkable that given the central importance of distribution functions to mathematical statisticians and probablists, that after more than 20 years of writing on the matter only one article has addressed HRX, or Scanlan's rule, within a more formal framework (Lambert and Subramanian 2014). Perhaps this is because Scanlan has not contributed to the mainstream sources most statisticians read. Whatever the reason, Lambert and Subramanian state but do not prove a result which is proven here in Theorem 2.1 and which holds in more generality. In particular, their statement (Lambert and Subramanian 2014, p. 568–569) applies to continuous random variables on $[0, \infty)$. So it does not apply to the most important case which Scanlan has employed in numerous settings as motivational framework for claiming HRX effects, the normal shift model. Stated here are conditions for continuous and discrete random variables and certain finite mixtures for which HRX holds. HRX is not, in general, a robust condition but can be very sensitive to model departures.

We start where Scanlan typically starts, with an empirical example. One of Scanlan's most readable shorter statements is his 2006 *Chance* article "Can We Actually Measure Health Disparities?" which compares poverty for both blacks and whites. Table 1 is based on U. S. Census 2004 poverty data (U. S. Census Bureau 2005). It is nearly identical to Scanlan's own Table 1, and differs from his largely in units and notation.

In Table 1 x is a percent poverty line; $x = 100$ is the 100% poverty line aggregate dollar amount defined by the U. S. Census Bureau; $x = 200$ and $x = 50$ would designate twice the poverty line income or half the income, respectively. In the following $F_B(x)$ replaces $F_D(x)$ and similarly $F_W(x)$ replaces $F_A(x)$. So $F_B(x)$ is the proportion of blacks below a given x, and similarly $F_W(x)$ for whites. Table 1 reveals that at all levels of x there are proportionally more blacks below x than whites. Second note that in Table 1, increases in x over the entire range of reported data (from 50 to 600) is associated with monotonic decreases in both SR and FR. This is HRX, which can be more easily seen by the graphed points in Fig. 1: As x decreases, both the empirical SR and FR increase. Consider those at the poverty line, $x = 100$, then $FR(100) = 2.29$ or proportionally more than twice as many blacks as whites fall at or below the poverty line. However the success ratio, $SR(100) = .84$, avoiding poverty, is nearer to one, and does not appear as bleak. The fact that FR and SR can display very different magnitudes of departure from one is one of the well know difficulties in deciding which risk ratio, FR or SR, on which to focus.

Table 1 Blacks and whites below percentage poverty line values x

x	$F_B(x)$	$F_W(x)$	$FR(x)$	$SR(x)$
600	.919	.795	1.16	.40
500	.869	.715	1.22	.46
400	.786	.605	1.30	.54
300	.662	.457	1.45	.62
250	.581	.373	1.56	.67
200	.488	.285	1.71	.72
185	.458	.258	1.78	.73
175	.437	.230	1.83	.74
150	.374	.191	1.96	.77
135	.333	.167	1.99	.80
130	.319	.156	2.04	.81
125	.310	.149	2.08	.81
100	.247	.108	2.29	.84
75	.179	.072	2.49	.88
50	.118	.044	2.68	.92

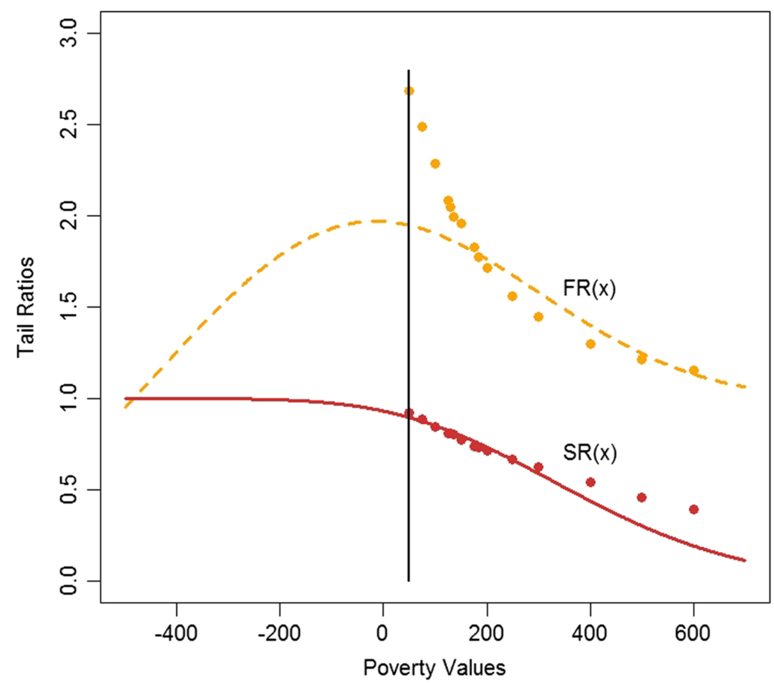

Fig. 1 Percent Poverty Table 1; points are empirical ratios; vertical is lower bound of 50; lines are fits under normality

It is important to recognize what Scanlan views as important. A naive glance at the $FR(100) = 2.29$ might prompt the conclusion that discrimination is the root cause of more than twice the proportion of blacks being below the poverty line than whites. Scanlan does not dispute this possibility. Rather, he argues that the property of HRX must be considered before a discrimination or bias conclusion is made. Thus, HRX, a purely statistical property of tail ratios may be the basis for the disparity, and not discrimination. Thus, HRX is a *possible* explanation that must be considered when disparities are interpreted.

Now suppose that both blacks and whites, through some effective social policy, experienced an additive shift such that those now at $x = 50$, or half the poverty line income, were now at $x = 100$, which can be viewed as effectively rescaling the x variable. While the overall proportions of both blacks and whites below the poverty line would both be substantially smaller, the corresponding FR would be much higher, going from 2.29 prior to the additive shift to $.118/.044 = 2.68$ afterwards. Hence, as Scanlan says "…the rarer the outcome, the greater the disparity in experiencing the outcome." Similarly, for the other ratio, "…the smaller the disparity in avoiding the outcome" namely the SR ratio associated with those avoiding poverty. The SR disparity is reduced from .84 to .92.

The change in these ratios, as x changes is the key critical empirical fact of Table 1 to note: As x increases, both FR and SR decrease. Scanlan argues that people (even professionals) often fail to appreciate how these ratios behave under varying values of x, at least in some settings. Without addressing what might be a plausible distribution of poverty rates, Scanlan has suggested that "any set of data reflecting more or less normal distributions of factors" will reveal HRX. (Scanlan 2006, p. 47–48). This claim will be considered below.

Early on, Scanlan considered risk ratios as they are defined above with the distribution of the advantaged group in the denominator. Consequently as Table 1 shows, both *FR* and *SR* decrease with increase in x. More recently, Scanlan has placed the larger of the tail areas in the numerator. So for example in Table 1, *SR* becomes for Scanlan $1/SR$, and consequently he now claims, in some of his writings, ..." the two relative differences tend to change in opposite direction ...(Scanlan 2014a, p. 330). Here "relative difference" is defined as the risk ratio minus one.

As suggested, Scanlan has addressed myriad settings where population disparities in outcomes occur, infant mortality, arrest frequencies, mortgage lending rates, school discipline rates and the like. The Scanlan signature model (SSM) which forms the conceptual basis for many of his empirical claims consists of two normal distributions with different means and equal variances, the normal shift model.

Let $n(\mu, \sigma)$ denote a normal density with mean μ and standard deviation σ. Denote the density for the disadvantaged group as $n_D(0, 1)$ and for the advantaged group as $n_A(.51, 1)$. Let Φ be the standard normal distribution function lower tail. Scanlan typically considers two decision values, $x = c = -1.13$ and $x = c = -.33$. Then $\Phi_D(-.33) = .37$ and $\Phi_D(-1.13) = .13$, $\Phi_A(-.33 - .51) = .20$, $\Phi_A(-1.13 - .51) = .05$, with the corresponding $FR(c)$, 1.85, and 2.55, and $SR(c)$ .787 and .915; these risk ratio values are shown in Fig. 2. Figure 2 also shows the two densities, the two $x = c$ cutting or decision values, and their associated risk ratios, as well as the trajectory of $FR(x)$ and $SR(x)$ over $-4 < x < 4$. Also shown is Scanlan's now favored upper tail ratio $1/SR(x)$. Note that $SR(x)$ and $FR(x)$ are decreasing functions. Note further their relation to the unity line at one.

What is often surprising is that, as the results of the empirical example in Table 1 indicated, both $SR(x)$ and $FR(x)$ show decreases with increases in x or alternatively,

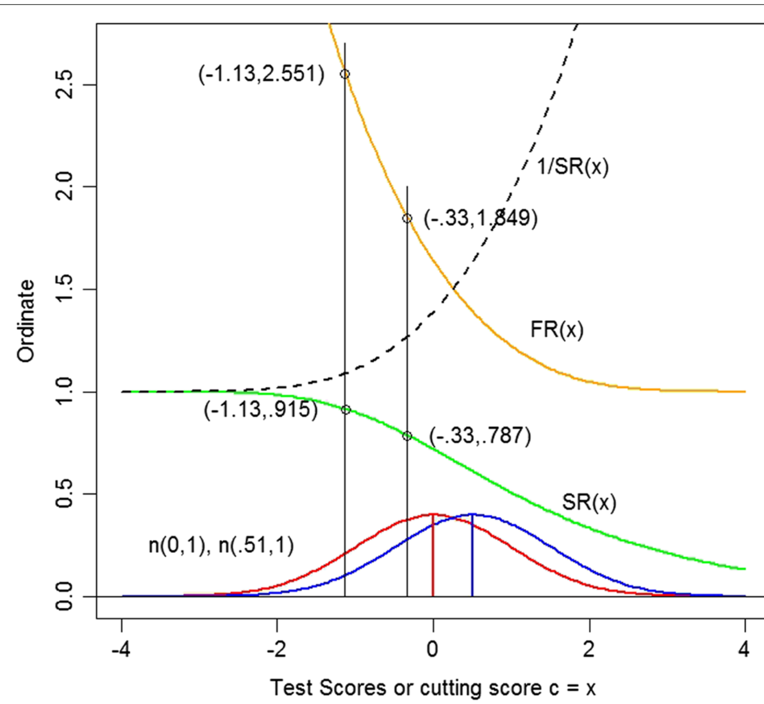

Fig. 2 Scanlan's signature normal shift model

increases with decreasing x, and in this case, over the entire range of x. Again, this is HRX. According to Scanlan: "In 2000, however, this pattern was virtually unknown to health researchers or anyone else..." (Scanlan 2014a, p. 328).

To understand HRX likely involves graphs, and grasping the trajectory of risk ratios $SR(x)$ and $FR(x)$ over x. Consequently, given the importance for Scanlan of this motivating case, graphs similar to Fig. 2 would seem natural candidates for display. It is thus puzzling that Scanlan apparently has never provided such a graph. He does provide poverty data risk ratio graphs in Scanlan (2006). No graphs appear in his most comprehensive publication (Scanlan 2014a). Two shifted normal densities appear in Scanlan (1994), but these do not illustrate HRX. He does *table* two, or sometime four decision c values and their corresponding numerical values (Scanlan 2014a, Table 1; Scanlan 2016b, Table 1, p. 416). In his presentation slide files empirical graphs of $FR(x)$ and $1/SR(x)$ appear, but they are not graphed against x as might be expected. Rather, these ratios are graphed against the failure rate $F_A(x)$. Furthermore, the $F_A(x)$ abscissa is numbered *backwards* reading from left to right .99, .98., ... .01 (e.g., Scanlan 2014b). It seems fair to say such graphs can certainly lead to discomfort if not confusion.

It is the equal variances case *only* that leads to HRX holding in the normal cases, over all x. See Example 4 below. In fact, seemingly small departures can lead to dramatic failures of HRX, as Fig. 3 shows, where $n_A(.5, 1)$, and $n_D(.4, 1.1)$. So caution needs to be exercised regarding claims about the effects of HRX in settings where the distributions are unknown. Still, HRX often seems to hold over regions of X which are of practical relevance for many empirical data settings. However, none of Scanlan's examples seems well modeled by SSM, his signature normal shift model. For example if the poverty data were regarded as being normal in distribution, then the values in Table 1 lead to mean and standard deviation estimates $n_B(222, 197), n_W(351, 233)$ with corresponding SR and FR

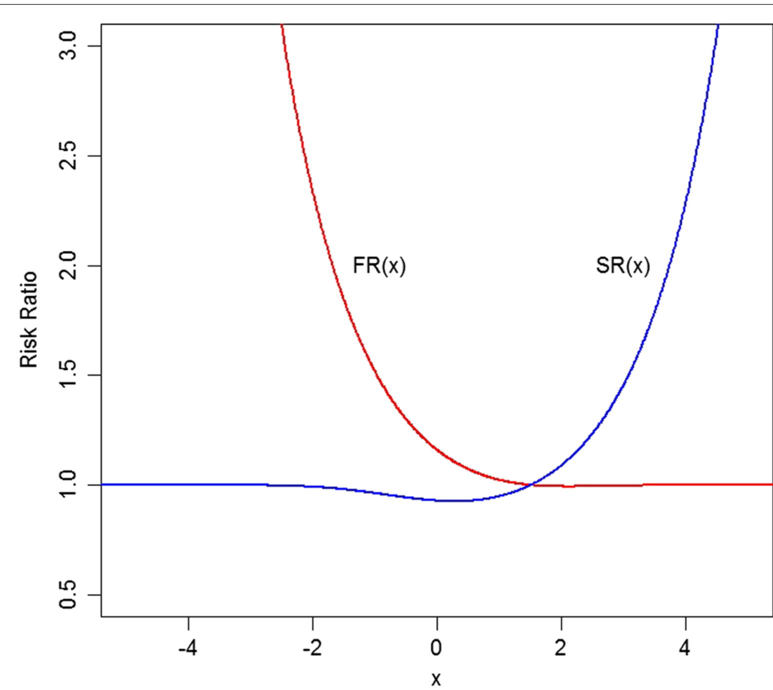

Fig. 3 $SR(x)$ and $FR(x)$ for $n_A(.5, 1), n_D(.4, 1.1)$

tail ratios in Fig. 1. Under normality HRX fails in general in this example when poverty values turn negative. However, HRX clearly holds over the realizable values of interest $50 \leq x \leq 600$ the range of Census Bureau data. Figure 1 also reveals that the empirical risk ratios (points) and model values of $SR(x)$ and $FR(x)$ (lines) constructed under normality are in poor agreement, signaling normality is a poor model for these poverty data.

Perhaps the setting where the effects of HRX and Scanlan's arguments should find the most saliency, are college admissions or similar selection settings. Here properties of the setting are reasonably well known. College officials have known control over the critical test scores required for acceptance, and the distribution of the test scores, such as SAT scores, are commonly assumed to be normal or roughly so. Different ethnic groups are known to often have very different SAT test score distributions, and these typically remain fairly stationary over the years. Consider African American and white student applicants; the mean difference for different portions of the SAT has remained essentially constant favoring whites for many years. The standard deviations have remained similar. From the National Center for Education Statistics (NCES) 2014, (or similarly, the College Board 2016) the reading SAT is for African Americans $n_B(429, 103)$ and for whites $n_W(531, 100)$. Figure 4 shows the risk ratios for these distributions for the SAT range of 400 to 700. HRX clearly holds here. There is no really satisfactory critical SAT that results in a success ratio near one so there is substantial disparity; $SR(600) = .20, SR(400) = .68$. With a acceptance score of 550 or above, $SR(550) = .28$, still a potentially socially troublesome value; the fail ratio, $FR(550) = 1.53$, so more African Americans are likely to be rejected than whites. However, this example may be unrealistic for some settings. Student applicants do self-select, and typically apply only to those colleges for which they have

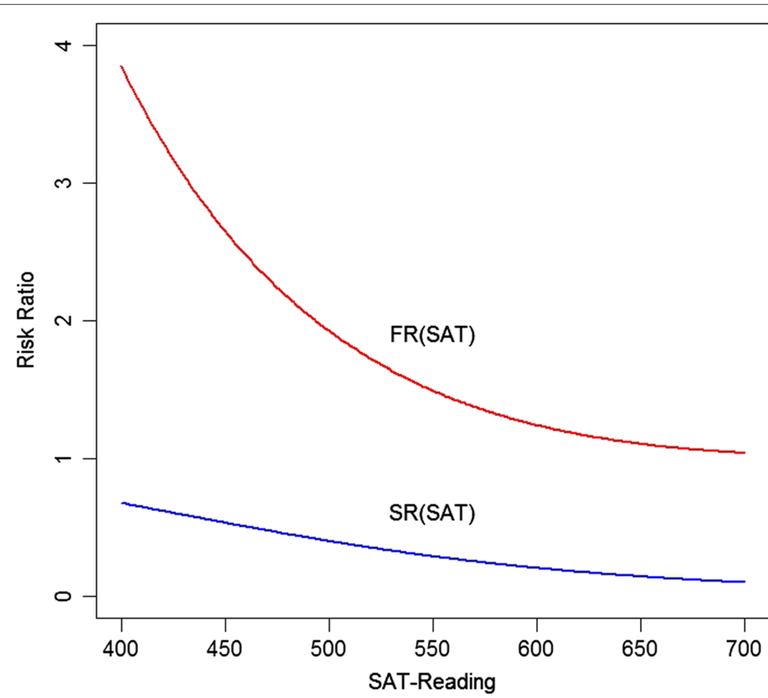

Fig. 4 $SR(SAT)$ and $FR(SAT)$ for $n_W(531, 100), n_B(429, 103)$

some reasonable expectation of being accepted. Consequently, the population distribution of college applicants for a given college may well be very different from the NCES data.

With the above as background, next consider conditions under which HRX holds, and conditions for HRX to hold for selected distributions. Some other related analytical results as presented as well.

2 Analytical results

2.1 Definitions and main results

Definition 1 *As above, let $F_D(x)$ and $F_A(x)$ be two cdfs with either probability density functions (pdfs) or probability mass functions (pmfs) $f_D(x)$ and $f_A(x)$ with support sets $\mathcal{S}_\mathcal{D}$ and $\mathcal{S}_A$, respectively. As defined at the outset above, $FR(x)$ and $SR(x)$ are failure and success ratios respectively. Assume both are non-increasing for $x \in \mathcal{S}_\mathcal{D} \cap \mathcal{S}_A$, with $FR(x) \geq 1$ and $SR(x) \leq 1$ then $F_D(x)$ and $F_A(x)$ are said to have the HRX property, or simply $F_D(x)$ and $F_A(x)$ are HRX.*

As noted above, Lambert and Subramanian (2014) consider without proof conditions under which continuous $F_D(x)$ and $F_A(x)$ are HRX, assuming $x \in [0, \infty)$. They relate HRX to a monotone likelihood ratio (MLR) property and stochastic ordering.

Definition 2 *$F_D(x)$ and $F_A(x)$ have the MLR property if $f_D(x)/f_A(x)$ is non-increasing for $x \in \mathcal{S}_\mathcal{D} \cap \mathcal{S}_A$.*

Definition 3 *$F(x)$ is called log-concave provided $\log f(x)$ is concave.*

Definition 4 *$F_D(x)$ and $F_A(x)$ form a location shift model when $F_A(x) = F_D(x-\theta), \theta > 0$, and $F_D(x)$ is a continuous distribution.*

The main result is stated in the following theorem which uses methods in Shaked and Shanthikumar (2007, Section 1.C.1).

Theorem 1 *If $F_D(x)$ and $F_A(x)$ have the MLR property, Definition 2, then they have property HRX of Definition 1. That is, MLR implies HRX.*

Proof Suppose $y_1, y_2 \in \mathcal{S}_D \cap \mathcal{S}_A$ and $y_1 \leq y_2$. Then $f_D(y_1)f_A(y_2) \geq f_D(y_2)f_A(y_1)$. Take $u_1, u_2 \in \mathcal{S}_D \cap \mathcal{S}_A$, and $u_1 \leq u_2$. Then

$$\int_{-\infty}^{u_1} dF_D(y_1) \int_{u_1}^{u_2} dF_A(y_2) \geq \int_{-\infty}^{u_1} dF_A(y_1) \int_{u_1}^{u_2} dF_D(y_2)$$

Expanding gives

$$F_D(u_1)\left[F_A(u_2) - F_A(u_1)\right] \geq F_A(u_1)\left[F_D(u_2) - F_D(u_1)\right]$$

and

$$FR(u_1) = \frac{F_D(u_1)}{F_A(u_1)} \geq FR(u_2) = \frac{F_D(u_2)}{F_A(u_2)}.$$

Hence, $FR(x)$ is non-increasing on $\mathcal{S}_D \cap \mathcal{S}_A$. A similar argument shows $SR(x)$ is also non-increasing on $\mathcal{S}_D \cap \mathcal{S}_A$. □

Wellner (2012) proved the following lemma:

Lemma 1 *The pdfs $f(x), f(x-\theta)$ have MLR property if and only if f is log-concave. Consequently a corollary to Theorem* 1 *is:*

Corollary 1 *Given a location shift model (Definition* 4*), if $F_D(x)$ is log-concave, then $F_D(x)$ and $F_A(x) = F_D(x-\theta), \theta > 0$, are HRX.*

Example 1 *Symmetric log-concave densities are often used in location shift models. These include the normal (noted above), as well as the uniform, logistic and Laplace.*

Example 2 *$F_D(x)$ and $F_A(x)$ need not be symmetric. This is the case for distributions with positive support. Assume the advantaged group mean exceeds the disadvantaged group mean, then the lemma can be used to establish HRX from Theorem* 1*. The following densities are log-concave: extreme value, gamma with shape parameter at least one, exponential, beta with both parameters at least one, and the Weibull with shape parameter at least one.*

The family of t distributions with finite degrees of freedom is not log-concave; hence by the lemma, t distributions do not satisfy the MLR property. This family includes the Cauchy density. The Pareto and lognormal are also not log-concave. We next consider the situation for which HRX holds for a subset of $\mathcal{S_D} \cap \mathcal{S_A}$.

Example 3 *Assuming the advantaged group mean exceeds the disadvantage group mean then it is straight forward to show the Poisson, negative binomial, both satisfy Definition* 2*, are MLR and are thus HRX. However, the binomial requires some care. Let $f_D(x) = \binom{n}{x} p_D^x (1-p_D)^{n-x}, x = 0, \ldots, n, f_A(x) = \binom{m}{x} p_A^x (1-p_A)^{m-x}, x = 0, \ldots, m$ with $p_A > p_D$, Note that $\mathcal{S_D} \cap \mathcal{S_A} = \{0, \ldots, \min(m,n)\}$. Define $r = p_D(1-p_A)/[(1-p_D)p_A]$ then $r < 1$, and*

$$\frac{f_D(x)}{f_A(x)} \geq \frac{f_D(x+1)}{f_A(x+1)} \iff x \leq (m-rn)/(1-r)$$

By considering cases $\min(m,n) = m$ or n, it follows the above expression holds if and only if $\min(m,n) - 1 \leq (m-rn)/(1-r)$. Hence, the binomial satisfies Definition 2*, and by Theorem* 1*, the HRX property holds. Discrete models can be used when the D and A populations must obtain a certain number of items correct to pass the test. If the criterion for passing is lowered to increase SR(x), FR(x) will unexpectedly increase, because of HRX.*

Mixtures provide flexible models, as will be illustrated below. Consider two component mixtures $F_D(x) = \lambda F_{D_1}(x) + (1-\lambda)F_{D_2}(x)$ and $F_A(x) = \alpha F_{A_1}(x) + (1-\alpha)F_{A_2}(x)$ with $0 < \lambda < 1, 0 < \alpha < 1$. The following theorem provides conditions under which $F_D(x)$ and $F_A(x)$ are HRX.

Theorem 2 *Given the above mixtures, suppose according to Definition* 2*, the following pairs have the MLR property $(f_{D_1}(x), f_{A_1}(x))$, $(f_{D_2}(x), f_{A_1}(x))$ and $(f_{A_1}(x), f_{A_2}(x))$, then the mixtures $f_D(x)$ and $f_A(x)$ are HRX.*

Proof

$$\frac{f_D(x)}{f_A(x)} = \frac{\lambda \frac{f_{D_1}(x)}{f_{A_1}(x)} + (1-\lambda)\frac{f_{D_2}(x)}{f_{A_1}(x)}}{\alpha + (1-\alpha)\frac{f_{A_2}(x)}{f_{A_1}(x)}}$$

The numerator is decreasing by the MLR property. The denominator is increasing because $\frac{f_{A_1}(x)}{f_{A_2}(x)}$ is decreasing. Hence, the ratio is decreasing and HRX follows from Theorem 1. □

A simple example of Theorem 2 occurs when the mixtures are defined by a log-concave location shift model with location parameters $\theta_{D_1} < \theta_{D_2} < \theta_{A_1} < \theta_{A_2}$. More complicated models are possible.

Recall the dilemma of HRX. Consider a criterion value c above which there is success. If c is lowered so that $SR(c)$ increases toward one, the unexpected consequence is that $FR(c)$ also increases. There is an alternative way to view this situation. Consider $P(D|F)$ as the proportion of D members in the failure pool and $P(D|S)$ as the proportion of D members in the success pool. Further, let $P(D)$ and $P(A)$ be the relative proportions in the $D \cup A$ population. We have the corresponding dilemma: When c is lowered so that $SR(c)$ increases, it also happens that $P(D|S)$ increases. Similarly, with increases in $FR(c)$ there are unexpected increases in $P(D|F)$; the proportion of the D population in the failure pool increases. This result is developed in the following theorem.

Theorem 3 *Suppose $F_D(x)$ and $F_A(x)$ satisfy Definition* 1 *and are HRX. Then $P(D|S)$ and $P(D|F)$ are non-decreasing functions of $SR(x)$ and $FR(x)$ respectively.*

Proof Let c be the criterion value for pass-fail. Note that $P(F|D) = F_D(c)$ and $P(F|S) = F_A(c)$. Using Bayes theorem:

$$P(D|F) = \frac{P(F|D)P(D)}{P(F|D)P(D) + P(F|A)P(A)}$$

Divide numerator and denominator by $P(F|A)$ then:

$$P(D|F) = \frac{FR(c)P(D)}{FR(c)P(D) + P(A)}.$$

This is a non-decreasing function of $FR(c)$. A similar argument shows $P(D|S)$ is a non-decreasing function of $SR(c)$. □

Suppose c is decreased. Under HRX both $SR(c)$ and $FR(c)$ increase. By Theorem 3 both $P(D|S)$ and $P(D|F)$ increase.

2.2 When HRX fails to hold

Example 4 *Consider the normal location-scale model and without loss of generality $F_A(x) = F_D\left(\frac{x-\theta}{\sigma}\right), \sigma > 0, \theta > 0$, with $F_D(x)$ the standard normal cdf. Then $\mathcal{S_D} \cap \mathcal{S_A} = (-\infty, \infty)$. The likelihood ratio $l(x)$ is*

$$l(x) = \exp\left[-2^{-1}x^2\right] / \exp\left[-\left(2\sigma^2\right)^{-1}\left(x-\theta^2\right)\right]$$

$l'(x) < 0$ when $1-\sigma^2 > 0$ and $x < \theta/(1-\sigma^2)$ or $1-\sigma^2 < 0$ and $x > \theta/(1-\sigma^2)$. The MLR property only holds for certain x. Figure 5 *shows that $FR(x)$ and $SR(x)$ are not decreasing functions of x on $(-\infty, \infty)$, however this example and others demonstrate there may be intervals for which HRX holds.*

Theorem 4 *Given continuous distributions and $FR(x), SR(x)$, and $l(x) = f_D(x)/f_A(x)$ the likelihood ratio, define $A = \{x : l(x) \leq FR(x)\}, B = \{x : SR(x) \leq l(x)\}, C = A \cap B$, then* *on the interval* (a, b) *where* $a = \inf C$ *and* $b = \sup C$*,* $FR(x)$ *and* $SR(x)$ *are non-increasing. Furthermore, when* $l(x)$ *is monotone decreasing on* (a, b) *then* a *and* b *are determined by*

$$SR(a) = l(a), FR(b) = l(b) \tag{1}$$

Proof The results follow from noting that the derivatives

$$SR'(x) \le 0 \iff SR(x) \le l(x) \text{ and } FR'(x) \le 0 \iff l(x) \le FR(x)$$

□

To apply Theorem 4 requires at a minimum the solution of Eq. 1. A more practical approach is to plot $FR(x)$ and $SR(x)$ and by inspection determine the HRX interval, noting whether the HRX interval contains the relevant critical values of c. As noted earlier, the t distribution is not log-concave, and hence the MLR property does not hold.

2.3 Data, inference, asymptotics

Suppose there are random samples from populations D and A, $D_1, D_2, \ldots, D_{n_D}$ and $A_1, A_2, \ldots, A_{n_A}$, and estimates $\widehat{SR}(c)$ and $\widehat{FR}(c)$ are obtained from the empirical cdfs $\hat{F}_D$ and $\hat{F}_A$. The samples are independent and $n_D\hat{F}_D(c)$ and $n_A\hat{F}_A(c)$ have binomial distributions. Suppose sample sizes increase at the same rate ($n_D/n_A \to k > 0$). Then asymptotic theory is based on the bivariate central limit theorem, along with Slutsky's theorem. The following results hold:

$$\sqrt{n_D + n_A} \begin{pmatrix} \widehat{FR}(c) - FR(c) \\ \widehat{SR}(c) - SR(c) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim BVN\left(\mathbf{0}, \mathbf{V}\right)$$

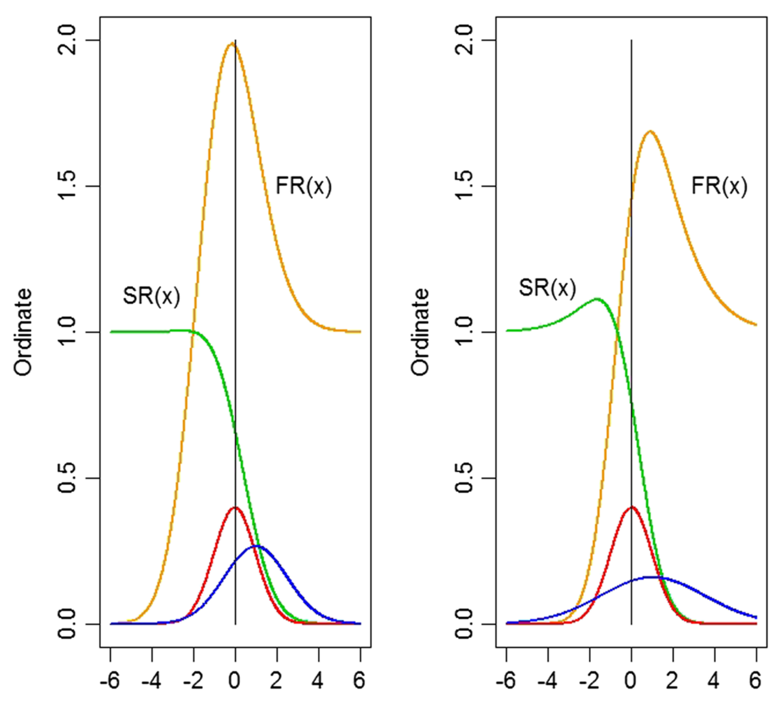

Fig. 5 Left: $n_D(0, 1)$, $n_A(1, 1.5)$; right: $n_D(0, 1)$, $n_A(1, 2.5)$

It is of interest to consider the variance, covariance and correlation between the risk ratio estimates $\widehat{SR}(c)$ and $\widehat{FR}(c)$ given Olkin's (1998) observation that "the risk ratio for dying is unrelated to the risk ratio for surviving..." and earlier Sheps (1958) commented that FR "has no predictable relation" to SR.

Consider a fixed c. The asymptotic covariance, variances and correlation are as follows:

$$\mathrm{cov}(\widehat{SR}(c), \widehat{FR}(c)) \approx -(1/n_A + 1/n_D) SR(c) \times FR(c)$$

and the variances are

$$\mathrm{var}(\widehat{SR}(c)) \approx SR(c)^2 \times \left(\frac{F_D(c)}{n_D(1 - F_D(c))} + \frac{F_A(c)}{n_A(1 - F_A(c))} \right)$$

and

$$\mathrm{var}(\widehat{FR}(c)) \approx FR(c)^2 \times \left(\frac{1 - F_D(c)}{n_D F_D(c)} + \frac{1 - F_A(c)}{n_A F_A(c)} \right).$$

The correlation is: $\rho\left(\widehat{SR}(c), \widehat{FR}(c)\right) \approx$

$$-\frac{\frac{1}{n_D} + \frac{1}{n_A}}{\sqrt{\left\{ \frac{F_D(c)}{n_D(1-F_D(c))} + \frac{F_A(c)}{n_A(1-F_A(c))} \right\} \left\{ \frac{(1-F_D(c))}{n_D F_D(c)} + \frac{(1-F_A(c))}{n_A F_A(c))} \right\}}}.$$

A simulation of 1000 $FR(c), SR(c)$ pairs, each based on sample size 1000, with $c = -.33$ under SSM, yields $\rho = -.91$, while the Pearson $r = -.87$. Clearly the risk ratios are correlated.

3 Discussion

Scanlan argues in his 2014 *Society* article "Race and Mortality Revisited" that since his 2000 "Race and Mortality" *Society* article that "almost nothing said" about health disparities "has had a sound statistical basis (Scanlan 2014a, p. 330)." This is quite an indictment. Without considering the merits of the statement, there seems little doubt that Scanlan's efforts have not elicited the attention they deserve; it is also clear that some writers have avoided if not ignored his work. Scanlan (2014a, p. 344) discusses at length the history of his interaction with the authors of "Commission Paper: Healthcare Disparities Measurement" (Weissmann et al. 2011) prior to its final release. It is an 84 page document and contains 125 references. None reference Scanlan. Penman-Aguilar et al. (2016), another health disparities summary document, contains 60 references. Scanlan does not appear. (He issued a sharp critique, Scanlan 2016b).

Scanlan's own writing can be hard to penetrate, as was noted at the outset, and his lack of notation, avoidance of graphs, especially in his published articles would appear to be clear hindrances to understanding. Simple graphs, such as Fig. 1, would appear to make HRX much more easily understandable. At least graphs should supplement tabled values, to facilitate understanding of HRX which certainly can be elusive.

Recall SSM denotes Scanlan's signature model and is displayed in Fig. 2. SSM is ubiquitous in its use in Scanlan's documents and is the motivation for invoking HRX under SSM in myriad settings. SSM is a restrictive model framework; it assumes there is a measurable variable, such as test scores, over which there is assumed to be a pair of normal or at least approximately normal distributions, with equal variances; it is assumed that the empirical data of focus are in rough correspondence to the SSM model. It further assumes there is some fixed cutting or decision score, so that the risk ratios are plausibly defined. Many test score settings seem appropriately so modeled by SSM. A serious problem arises when SSM is applied to settings which may lack any resemblance to this scenario, and for which there are little or no data, nor any statistical or substantive theory available for guidance.

An example concerns arrest warrants. Arrests are counts. One might model arrests as Poisson or binomial variables but they are clearly not normal variables nor is a shift model appropriate for count data. Scanlan (2016c) seems unperturbed. He reports that The Department of Justice filed suit against the city of Ferguson, Missouri claiming disparate impact because the city issued too many arrest warrants than they should have, given the city's 67% black population makeup. "What the DOJ fails to understand, however, is that reducing the number of citations and arrest warrants will tend to increase, not decrease, the proportion of African Americans make up (sic) of persons cited and arrested" (Scanlan 2016c). An abbreviated statement of SSM to motivate its application then follows. Yet there appear to be no data on *relative* arrest frequencies, nor any suggestion as to what distribution such frequencies might plausibly follow, should such data be available. The U. S. Department of Justice investigation of Ferguson (2015) is of little help. It provides only crude raw frequency or percentage data. There appears to be no data nor theory to defend the appropriateness of SSM. Furthermore, in city policing, the idea that numerous police officers in complex law enforcement settings would be expected to behave in a fashion roughly similar to the way a college admissions officer behaves in setting an SAT test selection criterion seems hopelessly unrealistic. (A more appropriate approach would seem to be to view police arrest decisions as random effects.) Furthermore, Scanlan's claim "The pattern is essentially universal" is certainly false; HRX is not robust, and can be very fragile as is illustrated by Figs. 3 and 5.

There are other settings for which the appropriateness of SSM appears questionable. School discipline is an example. More African Americans are often disciplined than whites, and discrimination is often claimed. Whether SSM is appropriate for such settings is never addressed. Recently, Scanlan (2017) considered racial differences in incarceration rates, and with similar SSM reasoning states that "Similar patterns will tend to appear when *any outcome* is increasingly restricted to those most susceptible to it—not in every instance of course, but a good deal of the time (p. 2)" (emphasis added).

This is not to say HRX is wrong as an explanation for settings where substantial disparities are found, only that there is insufficient evidence to suggest HRX, under SSM, provides a plausible and creditable motivating framework for some of Scanlan's settings of interest. Scanlan's goal is to influence public policy, often judicial policy. However too often his basis for claiming HRX is based on reasoning by analogy: If HRX under SSM holds in test score settings, it is appropriate, he reasons, for this same outcome to apply to a diverse range of settings which show no resemblance to SSM. Such reasoning is unlikely to sway policy makers. To invoke SSM in frameworks where SSM is unrealistic can only weaken the credibility of Scanlan's arguments in other arenas where his arguments have cogency, and that outcome would be unfortunate.

The analysis here leaves open important analytical questions that deserve attention. What has been given here are sufficient conditions to assure that HRX holds over the region of support for x and conditions for HRX to hold for certain discrete and continuous distributions. Whether necessary conditions can be given remains open. An important question which is discussed in the analytical section above, concerns the problem of specifying conditions under which HRX holds over regions of practical relevance. That HRX fails under normality when test scores turn negative is of no practical concern.

Dynamic graphs with sliders to adjust the model parameters can result in real-time displays of the risk ratios using Mathematica's `Manipulate` command (Wolfram Research Inc. 2017). These graphs are very helpful for exploring the behavior of $SR(x)$ and $FR(x)$. It would appear that, under normality, as long as the mean of the disadvantaged group is substantially smaller than the mean of the advantaged group, considerable variation in the variance parameters of each distribution can be tolerated for HRX to hold over non-negative x or the region of interest. Figure 4 provides a real illustrative example, and so of course does Fig. 1. Another set of issues are *statistical*. Some asymptotic results were given above for estimates of $SR(x)$ and $FR(x)$ and their correlation, but there are many open questions of interest. For example, given data, can an interval of x for which HRX holds be specified? This question relates to Theorem 4 and the discussion that follows that theorem. Can a test of the appropriateness of HRX be developed?

A critically important question is raised by Scanlan's concerns: How does one evaluate, in data, the relative importance of HRX given possible bias or discrimination? It seems highly unlikely, even if HRX were more widely understood, that matters of social policy would be altered substantially from current practice unless it were possible to separate the influence on risk ratios of HRX from the influence of discrimination or bias. Scanlan's HRX is an interesting property that deserves much wider attention.

Acknowledgements
There are no acknowledgements to be made at this time.

Author's contributions
Both authors contributed equally to the manuscript. HT wrote the manuscript, and did the computing; TPH provided the analytical work. Both authors read and approved the final manuscript.

Competing interests
The authors declare they have no competing interests. There was no funding associated with this manuscript. The manuscript has not been submitted elsewhere.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Pennsylvania State University, Department of Psychology, 121 Moore Bldg, University Park, PA 16802, USA.
[2]Pennsylvania State University, Department of Statistics, Thomas Bldg, University Park, PA 16802, USA.

Received: 5 April 2017 Accepted: 27 July 2017
Published online: 15 November 2017

References
College Board: Total Group Profile Report. (2016). https://reports.collegeboard.org/pdf/total-group-2016.pdf. Accessed 19 May 2017
Ferguson, United States Department of Justice Civil Rights Division: Investigation of the Ferguson Police Department. March 4, 2015 (2015). https://www.courts.mo.gov/file.jsp?id=95274. Accessed 18 May, 2017
Lambert, PJ, Subramanian, S: Disparities in socio-economic outcomes: some positive propositions and their normative implications. Soc. Choice Welf. **43**, 565–576 (2014)
National Center for Education Statistics: Fast Facts. SAT Scores. (2014). https://nces.ed.gov/fastfacts/display.asp?id=171. Accessed 18 May 2017
Olkin, I: Odds ratios revisited. Evid.-Based Med. **3**, 71 (1998)
Penman-Aguilar, A, Talih, M, Huang, D, Moonesinghe, R, Bouye, K, Beckles, G: Measurement of health disparities, health inequities, and social determinants of health to support the advancement of health equity. J Public Health Manag. Pract. **22**(1 Supp), S33–S42 (2016)
Scanlan, JP: Divining difference. Chance. **7**(4), 38–39, 45 (1994)
Scanlan, JP: Race and mortality. Society. **37**(2), 29–35 (2000)

Scanlan, JP: Can we actually measure health disparities? Chance. **19**(2), 47–51 (2006)
Scanlan, JP: Race and mortality revisited. Society. **51**, 328–346 (2014a)
Scanlan, JP: Rethinking the Measurement of Demographic Differences in Outcome Rates Methods Workshop, Maryland Population Research Center University of Maryland College Park, Maryland, October 14 (2014b). Available from www.jpscanlan.com
Scanlan, JP: Amicus brief submitted to United States District Court Eastern District of Missouri, Eastern Division. No. 4:16-cv-000180-CDP (2016a). Available from www.jpscanlan.com
Scanlan, JP: The mismeasure of health disparities. J. Public Health Manag. Pract. **22**, 415–419 (2016b)
Scanlan, JP: Things the DoJ doesn't know about racial disparities in Ferguson. The Hill, February 22 (2016c)
Scanlan, JP: Racial Impact Statement Laws in New Jersey and Elsewhere. Federalist Soc., March 20 (2017)
Shaked, M, Shanthikumar, JG: Stochastic Orders. Springer, New York (2007)
Sheps, MC: Shall we count the living or the dead? N. Engl. J. Med. **259**, 1210–1214 (1958)
Talih, M, Huang, DT: Measuring progress toward target attainment and the elimination of health disparities in Healthy People 2020. Healthy People Stat Notes. **27**, 1–20 (2016)
United States Census Bureau 2004: CPS Poverty Table: POV-01. The CPS 2005 Annual Social and Economic Supplement asked participants to report their poverty status for 2004 (2005). https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pov/pov-01.2004.html. Accessed 3 Feb 2017
Weissman, JS, Betancourt, JR, Green, AR, Meyer, GS, Tan-McGrory, A, Nudel, JD, Zeidman, JA, Carrillo, JE: Commissioned Paper: Healthcare Disparities Measurement. Disparities Solution Center of Massachusetts General Hospital (2011). https://mghdisparitiessolutions.files.wordpress.com/2015/10/dsc-commissioned-paper.pdf. Accessed 6 Aug 2017
Wellner, J: Log-concave distributions: definitions, properties, and consequences. Presentation, University of Paris-Diderot, January 23 (2012). http://www.proba.jussieu.fr/pageperso/picard/Paris-Jan-p1-small.pdf. Accessed 10 Mar 2017
Wolfram Research Inc: Mathematica, Version 11.1. Champaign, IL (2017)