Functional generalizations of Hoeffding's covariance lemma and a formula for Kendall's tau

Ambrose Lo
Department of Statistics and Actuarial Science, The University of Iowa
241 Schaeffer Hall, Iowa City, IA 52242-1409, USA

Abstract

This note extends Hoeffding's lemma from the covariance between two random variables to that between transformations of random variables, and provides clarification of several existing functional generalizations. In the same spirit as these results, an explicit integral formula of Kendall's tau for general, possibly discontinuous random variables is also determined.

Keywords: Covariance; Hoeffding's lemma; Lebesgue-Stieltjes integral; measure of concordance; Kendall's tau

1 Introduction

The notion of covariance as a simple reflection of the strength of the linear dependence between two random variables arises ubiquitously in probability, statistics, and various related areas. Among the multitude of covariance formulae in the literature, the one that is most intimately linked to the study of dependence structures and stochastic orders is the formula attributed to Hoeffding (1940). *Hoeffding's formula*, also interchangeably referred to as *Hoeffding's lemma* in the sequel, exhibits the covariance between any square-integrable random variables X and Y as an explicit integral comparison between their joint survival function and marginal survival functions, *i.e.*,

$$\mathrm{Cov}[X,Y] = \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X>x, Y>y] - \mathbb{P}[X>x]\mathbb{P}[Y>y]\right) \mathrm{d}x\mathrm{d}y. \tag{1.1}$$

This integral representation (1.1) separates the joint survival function of the random vector (X,Y), where information about the dependence structure of (X,Y) is captured, from the marginal survival functions.

Since its genesis in Hoeffding (1940), there have been attempts in the literature to extend Hoeffding's lemma to transformations of random variables, but most have reached limited success. Under the assumption of absolutely continuous distributions, Mardia (1967) derived $\mathrm{Cov}(X^r, Y^s)$ for $r \geq 1$ and $s \geq 1$ in conjunction with some special contingency-type bivariate distributions. Building on the prototypical higher-order generalization of Mardia (1967), Sen (1994) mentioned, without proof, the adaptation of Hoeffding's formula to strictly monotone transformations of random variables. The generalized formula of Sen (1994) reads (see Equation (2.4) therein)

$$\mathrm{Cov}[f(X),g(Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X>x, Y>y] - \mathbb{P}[X>x]\mathbb{P}[Y>y]\right) \mathrm{d}f(x)\mathrm{d}g(y), \tag{1.2}$$

where f and g are strictly monotone functions, and the integrals are presumably interpreted in the Lebesgue-Stieltjes sense. However, as Section 2 of this article shows, Equation (1.2), albeit its simplicity and resemblance to (1.1), is generally flawed due to the possible discontinuities of the functions f and g, and that Lebesgue-Stieltjes integrals are traditionally defined only with respect to left-continuous or right-continuous non-decreasing integrators. More recently, Cuadras (2002, 2015) established the validity of (1.2) for continuous random variables and transformations of bounded variation. In brief, many of these functional extensions suffer from a faulty formulation or the imposition of technical continuity assumptions on the underlying random variables or the transformations, undermining the generality of the intended results and their applicability to real problems.

Motivated by the profound impact of Hoeffding's lemma, but limited generality of its existing functional extensions, this note aims to derive rigorously a series of generalizations of Hoeffding's lemma to the covariance between general transformations of general random variables, thereby correcting Sen (1994)'s conjecture and extending Cuadras (2002)'s formula. As a stepping stone, we first formulate in Section 2 Hoeffding's lemma for monotone transformations of random variables. Apparently distinct from Sen (1994)'s conjectured equation, the generalized covariance formula in this version turns out to comprise four Lebesgue-Stieltjes integrals with different integrands and integrators. With the first-step extension to monotone functions at hand, in Section 3 we further extend Hoeffding's lemma to a wider class of non-monotone functions, including absolutely continuous functions and functions of bounded variation. Finally, Section 4 shares the same spirit as Sections 2 and 3 and presents a closed-form formula for Kendall's tau for general, possibly discontinuous random variables.

Throughout this paper, all random variables are defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We do not assume a priori the continuity or strict monotonicity of their distribution functions. It is tacitly assumed, however, that all integrals and covariances exist and are finite.

2 Monotone functions

Before formulating our main results, it is instructive to review the deceptively simple notion of Lebesgue-Stieltjes integral. Recall that one can associate a given non-decreasing right-continuous real function f with a measure μ_f on $\mathbb{R}$ defined by

$$\mu_f\left((a, b]\right) = f(b) - f(a) \quad \text{for } a \leq b, \tag{2.1}$$

which extends uniquely to the Borel σ-algebra on $\mathbb{R}$. The right-continuity of f is crucial to establishing the countable additivity of μ_f. An analogous identification can be made if f is non-decreasing and left-continuous via the formula

$$\mu_f\left([a, b)\right) = f(b) - f(a) \quad \text{for } a \leq b. \tag{2.2}$$

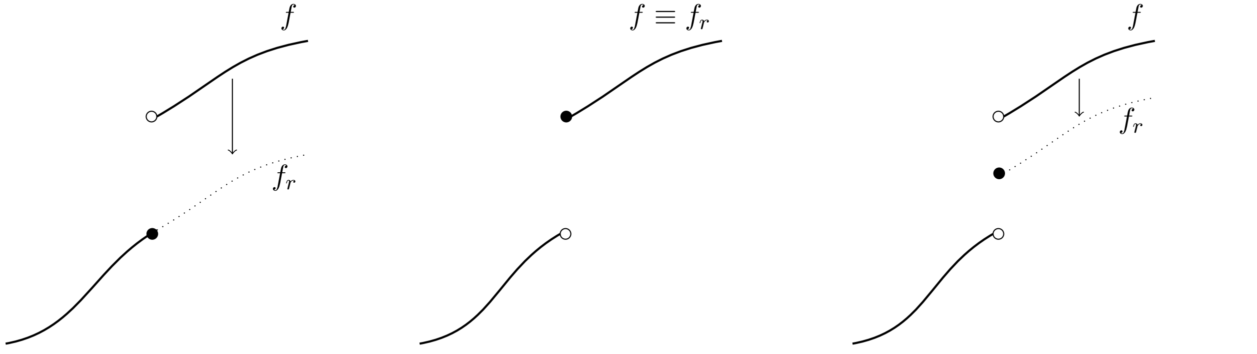
Notice that in (2.1), the interval in the argument of μ_f is open on the left, but closed on the right, whereas in (2.2) the interval is closed on the left, but open on the right. Such subtle differences will play a crucial role in the functional generalizations of this paper. For later purposes, we also point out that the measures defined in (2.1) and (2.2) are σ-finite (see, e.g., Theorem 2.8.1 of Leadbetter et al. (2014)), which is essential for Fubini's theorem to be applicable.

Given a non-decreasing left- or right-continuous function f and its associated measure μ_f, we define the *Lebesgue-Stieltjes integral* of a Borel measurable function ϕ with respect to f by

$$\int_{\mathbb{R}} \phi(x)\, \mathrm{d}f(x) := \int_{\mathbb{R}} \phi(x)\, \mathrm{d}\mu_f(x),$$

provided that the Lebesgue integral on the right-hand side exists.

To set forth our functional generalizations of Hoeffding's lemma, we need a technical lemma which suggests how Lebesgue-Stieltjes integrals with respect to a general (not necessarily left-continuous or right-continuous) monotone function can be defined unambiguously. Without loss of generality, we consider non-decreasing functions. Below we denote by $f(x_+)$ and $f(x_-)$ the right-hand and left-hand limits of a function f at x.



(a) Right-discontinuous point

(b) Left-discontinuous point

(c) Right- and left-discontinuous point

Figure 2.1: The construction of f_r in the proof of Lemma 2.1 in the case of a single discontinuity point. The line in bold represents the original function f.

Lemma 2.1. *Let f be a non-decreasing real function. There exist a non-decreasing right-continuous function f_r and a non-decreasing left-continuous function f_l such that $f = f_r + f_l$.*

Proof. For any real x, the function f must be either continuous at x (if $f(x_-) = f(x) = f(x_+)$) or have jump(s) arising from the left-discontinuity (if $f(x_-) < f(x)$) and/or right-discontinuity (if $f(x) < f(x_+)$) at x. Therefore, we can uniquely represent f as

$$f(x) = f_c(x) + \sum_i a_i 1_{[c_i,\infty)}(x) + \sum_j b_j 1_{(d_j,\infty)}(x), \quad \text{for all } x \in \mathbb{R}, \tag{2.3}$$

for some real c_i's which are the left-discontinuous points of f with a jump size of a_i (≥ 0), some real d_j's which are the right-discontinuous points of f with a jump size of b_j (≥ 0), with the two sums (empty sums are defined by convention to be zero) taken over the at most countable set of discontinuity of f, and $f_c(x) := f(x) - \sum_i a_i 1_{[c_i,\infty)}(x) - \sum_j b_j 1_{(d_j,\infty)}(x)$, which, by construction, is a continuous function. Defining

$$f_r(x) := f_c(x) + \sum_i a_i 1_{[c_i,\infty)}(x) \quad \text{and} \quad f_l(x) := \sum_j b_j 1_{(d_j,\infty)}(x), \tag{2.4}$$

which are a non-decreasing right-continuous function and a non-decreasing left-continuous function respectively, we immediately obtain $f = f_r + f_l$. □

Remark 2.2. (i) The constructions of f_r and f_l in (2.4) are motivated from the proof of Theorem 7 of Dhaene et al. (2012) in the context of distortion risk measures. Here the function f_l captures successively the magnitude of each right-discontinuous jump of f, and f_r is a modification of f by decrementing the graph of f strictly after each right-discontinuous jump by precisely the size of that jump, restoring right continuity (see Figure 2.1).

(ii) The Lebesgue-Stieltjes integral of a Borel measurable function h with respect to f can be defined as

$$\int_{\mathbb{R}} h(x)\,\mathrm{d}f(x) := \int_{\mathbb{R}} h(x)\,\mathrm{d}f_r(x) + \int_{\mathbb{R}} h(x)\,\mathrm{d}f_l(x), \tag{2.5}$$

where the two Lebesgue-Stieltjes integrals on the right-hand side of (2.5) are well-defined. This definition will be of use in Section 3. It is easy to see that this integral does not depend on the particular right-continuous and left-continuous components of f. More precisely, choosing

$$\tilde{f}_r(x) := \sum_i a_i 1_{[c_i,\infty)}(x) \quad \text{and} \quad \tilde{f}_l(x) = f_c(x) + \sum_j b_j 1_{(d_j,\infty)}(x) \tag{2.6}$$

for $f = \tilde{f}_r + \tilde{f}_l$ results in the same Lebesgue-Stieltjes integral with respect to f.

The following lemma providing integral representations for differences between functional values of non-decreasing functions will also prove handy.

Lemma 2.3. *Let h be a non-decreasing real function, and x_1 and x_2 be any real numbers.*

(a) If h is right-continuous, then

$$h(x_1) - h(x_2) = \int_{\mathbb{R}} \big(1_{[t,\infty)}(x_1) - 1_{[t,\infty)}(x_2)\big)\, \mathrm{d}h(t).$$

(b) If h is left-continuous, then

$$h(x_1) - h(x_2) = \int_{\mathbb{R}} \big(1_{(t,\infty)}(x_1) - 1_{(t,\infty)}(x_2)\big)\, \mathrm{d}h(t).$$

Proof. We only prove (a) because the proof of (b) is similar. Distinguishing between whether $x_2 < x_1$ or $x_1 < x_2$, we have

$$\begin{aligned} h(x_1) - h(x_2) &= 1_{(-\infty,x_1)}(x_2)[h(x_1) - h(x_2)] - 1_{(-\infty,x_2)}(x_1)[h(x_2) - h(x_1)] \\ &= 1_{(-\infty,x_1)}(x_2) \int_{\mathbb{R}} 1_{(x_2,x_1]}(t)\, \mathrm{d}h(t) - 1_{(-\infty,x_2)}(x_1) \int_{\mathbb{R}} 1_{(x_1,x_2]}(t)\, \mathrm{d}h(t) \\ &= 1_{(-\infty,x_1)}(x_2) \int_{\mathbb{R}} \big(1_{(-\infty,x_1]}(t) - 1_{(-\infty,x_2]}(t)\big)\, \mathrm{d}h(t) \\ &\qquad - 1_{(-\infty,x_2)}(x_1) \int_{\mathbb{R}} \big(1_{(-\infty,x_2]}(t) - 1_{(-\infty,x_1]}(t)\big)\, \mathrm{d}h(t) \\ &= \int_{\mathbb{R}} \big(1_{(-\infty,x_1]}(t) - 1_{(-\infty,x_2]}(t)\big)\, \mathrm{d}h(t) \\ &= \int_{\mathbb{R}} \big(1_{[t,\infty)}(x_1) - 1_{[t,\infty)}(x_2)\big)\, \mathrm{d}h(t), \end{aligned}$$

where the second equality follows from (2.1). □

We are now in a position to present the first generalized Hoeffding's lemma by modifying the coupling technique used to prove (1.1) (see Property 1.6.13 on page 28 of Denuit et al. (2005) amongst others for a typical proof of Hoeffding's lemma). The resulting formula comprises four Lebesgue-Stieltjes integrals, all of which possess distinct integrands and integrators. The integral representation is in stark contrast to (1.2), which fails to account for the possible discontinuities of the two monotone transformations.

Theorem 2.4. *Let X and Y be random variables, and f and g be non-decreasing real functions such that*

$$\mathbb{E}[|f(X)|] < \infty, \quad \mathbb{E}[|g(Y)|] < \infty \quad \textit{and} \quad \mathbb{E}[|f(X)g(Y)|] < \infty. \tag{2.7}$$

Then

$$\begin{aligned}
\mathrm{Cov}[f(X),g(Y)] \;=\;& \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X\geq x, Y\geq y]-\mathbb{P}[X\geq x]\mathbb{P}[Y\geq y]\right)\mathrm{d}f_r(x)\mathrm{d}g_r(y)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X\geq x, Y> y]-\mathbb{P}[X\geq x]\mathbb{P}[Y> y]\right)\mathrm{d}f_r(x)\mathrm{d}g_l(y)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X> x, Y\geq y]-\mathbb{P}[X> x]\mathbb{P}[Y\geq y]\right)\mathrm{d}f_l(x)\mathrm{d}g_r(y)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X> x, Y> y]-\mathbb{P}[X> x]\mathbb{P}[Y> y]\right)\mathrm{d}f_l(x)\mathrm{d}g_l(y) \qquad (2.8)
\end{aligned}$$

for any non-decreasing right-continuous functions f_r, g_r and non-decreasing left-continuous functions f_l, g_l such that $f = f_r + f_l$ and $g = g_r + g_l$.

Proof. Using a standard extension argument (see, e.g., page 111 of Kallenberg (2002)), we may assume without loss of generality that the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supports a random vector (X', Y') which is an i.i.d. copy of (X, Y). Then

$$\begin{aligned}
2\mathrm{Cov}[f(X),g(Y)] &= 2\mathbb{E}\left[\left(f(X)-\mathbb{E}[f(X)]\right)\left(g(Y)-\mathbb{E}[g(Y)]\right)\right]\\
&= \mathbb{E}\left[\big(f(X)-f(X')\big)\big(g(Y)-g(Y')\big)\right]. \qquad (2.9)
\end{aligned}$$

By Lemma 2.1, we fix any non-decreasing right-continuous functions f_r, g_r, and non-decreasing left-continuous functions f_l, g_l satisfying $f = f_r + f_l$ and $g = g_r + g_l$. Rewriting the two differences $f(X) - f(X')$ and $g(Y) - g(Y')$ in (2.9) as appropriate Lebesgue-Stieltjes integrals by virtue of Lemma 2.3 yields

$$\begin{aligned}
&2\mathrm{Cov}[f(X),g(Y)]\\
=\;& \mathbb{E}\left[\big(f_r(X)-f_r(X')\big)\big(g_r(Y)-g_r(Y')\big)\right]+\mathbb{E}\left[\big(f_r(X)-f_r(X')\big)\big(g_l(Y)-g_l(Y')\big)\right]\\
&+\mathbb{E}\left[\big(f_l(X)-f_l(X')\big)\big(g_r(Y)-g_r(Y')\big)\right]+\mathbb{E}\left[\big(f_l(X)-f_l(X')\big)\big(g_l(Y)-g_l(Y')\big)\right]\\
=\;& \mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{[x,\infty)}(X)-1_{[x,\infty)}(X'))\,\mathrm{d}f_r(x)\right)\left(\int_{\mathbb{R}}(1_{[y,\infty)}(Y)-1_{[y,\infty)}(Y'))\,\mathrm{d}g_r(y)\right)\right]\\
&+\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{[x,\infty)}(X)-1_{[x,\infty)}(X'))\,\mathrm{d}f_r(x)\right)\left(\int_{\mathbb{R}}(1_{(y,\infty)}(Y)-1_{(y,\infty)}(Y'))\,\mathrm{d}g_l(y)\right)\right]\\
&+\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{(x,\infty)}(X)-1_{(x,\infty)}(X'))\,\mathrm{d}f_l(x)\right)\left(\int_{\mathbb{R}}(1_{[y,\infty)}(Y)-1_{[y,\infty)}(Y'))\,\mathrm{d}g_r(y)\right)\right]\\
&+\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{(x,\infty)}(X)-1_{(x,\infty)}(X'))\,\mathrm{d}f_l(x)\right)\left(\int_{\mathbb{R}}(1_{(y,\infty)}(Y)-1_{(y,\infty)}(Y'))\,\mathrm{d}g_l(y)\right)\right]. \qquad (2.10)
\end{aligned}$$

We claim that these four expectations can be simplified into

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{[x,\infty)}(X)-1_{[x,\infty)}(X'))\,\mathrm{d}f_r(x)\right)\left(\int_{\mathbb{R}}(1_{[y,\infty)}(Y)-1_{[y,\infty)}(Y'))\,\mathrm{d}g_r(y)\right)\right]\\
&\qquad = 2\int_{\mathbb{R}}\int_{\mathbb{R}}\left(\mathbb{P}[X\geq x, Y\geq y]-\mathbb{P}[X\geq x]\mathbb{P}[Y\geq y]\right)\mathrm{d}f_r(x)\mathrm{d}g_r(y), \qquad (2.11)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{[x,\infty)}(X)-1_{[x,\infty)}(X'))\,\mathrm{d}f_r(x)\right)\left(\int_{\mathbb{R}}(1_{(y,\infty)}(Y)-1_{(y,\infty)}(Y'))\,\mathrm{d}g_l(y)\right)\right]\\
&\qquad = 2\int_{\mathbb{R}}\int_{\mathbb{R}}\left(\mathbb{P}[X\geq x, Y> y]-\mathbb{P}[X\geq x]\mathbb{P}[Y> y]\right)\mathrm{d}f_r(x)\mathrm{d}g_l(y), \qquad (2.12)
\end{aligned}$$

$$\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{(x,\infty)}(X)-1_{(x,\infty)}(X'))\,\mathrm{d}f_l(x)\right)\left(\int_{\mathbb{R}}(1_{[y,\infty)}(Y)-1_{[y,\infty)}(Y'))\,\mathrm{d}g_r(y)\right)\right]$$
$$=2\int_{\mathbb{R}}\int_{\mathbb{R}}(\mathbb{P}[X>x,Y\geq y]-\mathbb{P}[X>x]\mathbb{P}[Y\geq y])\,\mathrm{d}f_l(x)\mathrm{d}g_r(y), \qquad (2.13)$$

$$\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{(x,\infty)}(X)-1_{(x,\infty)}(X'))\,\mathrm{d}f_l(x)\right)\left(\int_{\mathbb{R}}(1_{(y,\infty)}(Y)-1_{(y,\infty)}(Y'))\,\mathrm{d}g_l(y)\right)\right]$$
$$=2\int_{\mathbb{R}}\int_{\mathbb{R}}(\mathbb{P}[X>x,Y>y]-\mathbb{P}[X>x]\mathbb{P}[Y>y])\,\mathrm{d}f_l(x)\mathrm{d}g_l(y), \qquad (2.14)$$

Since the techniques to derive (2.11)–(2.14) are highly similar, for the purpose of illustration we only show (2.13). To this end, we multiply the two integrals, interchange the order of integration by Fubini's theorem (justified by the integrability conditions given in (2.7)) and using the fact that (X,Y) and (X',Y') are i.i.d., yielding

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{\mathbb{R}}(1_{(x,\infty)}(X)-1_{(x,\infty)}(X'))\,\mathrm{d}f_l(x)\right)\left(\int_{\mathbb{R}}(1_{[y,\infty)}(Y)-1_{[y,\infty)}(Y'))\,\mathrm{d}g_r(y)\right)\right]\\
=\;&\int_{\Omega}\int_{\mathbb{R}}\int_{\mathbb{R}}\big(1_{\{X>x,Y\geq y\}}(\omega)-1_{\{X>x,Y'\geq y\}}(\omega)\\
&\qquad -1_{\{X'>x,Y\geq y\}}(\omega)+1_{\{X'>x,Y'\geq y\}}(\omega)\big)\,\mathrm{d}f_l(x)\mathrm{d}g_r(y)\mathrm{d}\mathbb{P}(\omega)\\
=\;&\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\Omega}\big(1_{\{X>x,Y\geq y\}}(\omega)-1_{\{X>x,Y'\geq y\}}(\omega)\\
&\qquad -1_{\{X'>x,Y\geq y\}}(\omega)+1_{\{X'>x,Y'\geq y\}}(\omega)\big)\,\mathrm{d}\mathbb{P}(\omega)\mathrm{d}f_l(x)\mathrm{d}g_r(y)\\
=\;&2\int_{\mathbb{R}}\int_{\mathbb{R}}(\mathbb{P}[X>x,Y\geq y]-\mathbb{P}[X>x]\mathbb{P}[Y\geq y])\,\mathrm{d}f_l(x)\mathrm{d}g_r(y).
\end{aligned}$$

Finally, inserting the expressions in (2.11)–(2.14) into (2.10) leads to the desired integral representation of $\mathrm{Cov}[f(X),g(Y)]$. □

Several cautionary remarks concerning the proof of Theorem 2.4 are in order.

Remark 2.5. (i) Due to the subtle differences between how the induced measures of right-continuous and left-continuous non-decreasing functions are defined in (2.1) and (2.2), the integrands of the corresponding Lebesgue-Stieltjes integrals differ. Specifically, strict inequalities go hand in hand with left-continuous integrators whereas weak inequalities and right-continuous integrators hang together.

(ii) It is seen from the proof of Theorem 2.4 that the validity of (2.8) is unaffected by which left-continuous and right-continuous components of f and g are selected.

(iii) One may be tempted to argue *erroneously* that because the functions

$$\begin{aligned}
(x,y) &\mapsto \mathbb{P}[X\geq x,Y\geq y]-\mathbb{P}[X\geq x]\mathbb{P}[Y\geq y]\\
(x,y) &\mapsto \mathbb{P}[X\geq x,Y>y]-\mathbb{P}[X\geq x]\mathbb{P}[Y>y]\\
(x,y) &\mapsto \mathbb{P}[X>x,Y\geq y]-\mathbb{P}[X>x]\mathbb{P}[Y\geq y]\\
(x,y) &\mapsto \mathbb{P}[X>x,Y>y]-\mathbb{P}[X>x]\mathbb{P}[Y>y]
\end{aligned}$$

are equal almost everywhere, it appeared valid to replace the four integrands in (2.8) by the same integrand and to use (2.5) to simplify (2.8) as a single integral:

$$\mathrm{Cov}[f(X),g(Y)]=\int_{\mathbb{R}}\int_{\mathbb{R}}(\mathbb{P}[X>x,Y>y]-\mathbb{P}[X>x]\mathbb{P}[Y>y])\,\mathrm{d}f(x)\mathrm{d}g(y),$$

which is Equation (2.4) of Sen (1994). The fact is that the above four bivariate functions are almost everywhere equal only with respect to the Lebesgue measure on $\mathbb{R}^2$, but the induced measures of the left- and right-continuous components of f and g need not be absolutely continuous with respect to the Lebesgue measure.

The following simple example refutes the general validity of Sen (1994)'s conjectured equation and shows that

$$\mathrm{Cov}[f(X), g(Y)] \neq \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y])\, \mathrm{d}f(x)\mathrm{d}g(y)$$

in general even when f and g are left-continuous or right-continuous non-decreasing functions.

Example 2.6. For any fixed real c and d, we take $f(x) = 1_{(c,\infty)}(x)$ and $g(y) = 1_{[d,\infty)}(y)$, which are left-continuous and right-continuous non-decreasing functions respectively. Following (2.4), we have $f_r = g_l = 0$, $f_l = f$ and $g_r = g$ with the induced measures of f and g being δ_c and δ_d, the Dirac measures at c and d respectively. Applying Theorem 2.4, we have

$$\begin{aligned}\mathrm{Cov}\left[1_{(c,\infty)}(X), 1_{[d,\infty)}(Y)\right] &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y \geq y] - \mathbb{P}[X > x]\mathbb{P}[Y \geq y])\, \mathrm{d}\delta_c(x)\mathrm{d}\delta_d(y) \\ &= \mathbb{P}[X > c, Y \geq d] - \mathbb{P}[X > c]\mathbb{P}[Y \geq d],\end{aligned}$$

which agrees with the elementary evaluation

$$\begin{aligned}\mathrm{Cov}\left[1_{(c,\infty)}(X), 1_{[d,\infty)}(Y)\right] &= \mathbb{E}[1_{(c,\infty)\times[d,\infty)}(X, Y)] - \mathbb{E}\left[1_{(c,\infty)}(X)\right]\mathbb{E}\left[1_{[d,\infty)}(Y)\right] \\ &= \mathbb{P}[X > c, Y \geq d] - \mathbb{P}[X > c]\mathbb{P}[Y \geq d].\end{aligned}$$

Incidentally, note that in general (e.g. if the distribution of (X, Y) has an atom at (c, d))

$$\begin{aligned}\mathrm{Cov}[f(X), g(Y)] &\neq \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y])\, \mathrm{d}f(x)\mathrm{d}g(y) \\ &= \mathbb{P}[X > c, Y > d] - \mathbb{P}[X > c]\mathbb{P}[Y > d].\end{aligned}$$

The generalized Hoeffding's formula in Theorem 2.4 for general monotone transformations consists of four distinct integrals and may look formidable. Under appropriate continuity assumptions on either the monotone transformations (which apply to Mardia (1967) because $f(x) = x^r$ and $g(y) = y^s$ are continuous) or the joint distribution function of the two random variables (as in Cuadras (2002)), considerable simplification arises and the four-term formula reduces to Sen (1994)'s formula, namely (1.2).

Corollary 2.7. *Let X and Y be random variables, and f and g be non-decreasing real functions such that* (2.7) *holds.*

(a) If both f and g are right-continuous, then

$$\mathrm{Cov}[f(X), g(Y)] = \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X \geq x, Y \geq y] - \mathbb{P}[X \geq x]\mathbb{P}[Y \geq y])\, \mathrm{d}f(x)\mathrm{d}g(y).$$

(b) If f is right-continuous and g is left-continuous, then

$$\mathrm{Cov}[f(X), g(Y)] = \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X \geq x, Y > y] - \mathbb{P}[X \geq x]\mathbb{P}[Y > y])\, \mathrm{d}f(x)\mathrm{d}g(y).$$

(c) If f is left-continuous and g is right-continuous, then

$$\mathrm{Cov}[f(X),g(Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X > x, Y \geq y] - \mathbb{P}[X > x]\mathbb{P}[Y \geq y]\right) \mathrm{d}f(x)\mathrm{d}g(y).$$

(d) If both f and g are left-continuous, then

$$\mathrm{Cov}[f(X),g(Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]\right) \mathrm{d}f(x)\mathrm{d}g(y).$$

(e) If f and g are continuous, or the joint distribution function of (X,Y) is continuous, then each of the identities under (a) to (d) holds.

3 Absolutely continuous functions

We now enhance the applicability of Hoeffding's lemma by extending its coverage to the covariance between non-monotone functions of random variables. The key is to decompose the transformations f and g into differences of appropriate non-decreasing functions, to which Theorem 2.4 can be applied. We first examine the important case of absolutely continuous functions, for which the simple double integral representation of covariance can be transformed into ordinary Lebesgue integrals. Our functional generalization of Hoeffding's lemma is concluded by a discussion on how general discontinuous and non-monotone transformations can be dealt with.

By definition, a real function f is said to be *absolutely continuous* if for every $\epsilon > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^{n} |f(b_i) - f(a_i)| < \epsilon$ whenever $\{(a_i, b_i)\}_{i=1}^{n}$ is a finite disjoint collection of open intervals in $\mathbb{R}$ with $\sum_{i=1}^{n}(b_i - a_i) < \delta$. To define the Lebesgue-Stieltjes integral with respect to a general absolutely continuous function f, we first write $f = f_1 - f_2$ for some non-decreasing absolutely continuous functions f_1 and f_2 (see, e.g., Theorem 39.11 of Aliprantis and Burkinshaw (1998)). Then the Lebesgue-Stieltjes integral of a Borel measurable function h with respect to f is defined as

$$\int_{\mathbb{R}} h(x)\,\mathrm{d}f(x) := \int_{\mathbb{R}} h(x)\,\mathrm{d}f_1(x) - \int_{\mathbb{R}} h(x)\,\mathrm{d}f_2(x), \tag{3.1}$$

provided that the two Lebesgue-Stieltjes integrals on the right-hand side exist, and the difference makes sense. By standard measure-theoretic techniques, it can be shown that such a definition does not depend on the particular f_1 and f_2 chosen to represent f (see page 379 of Aliprantis and Burkinshaw (1998)).

Theorem 3.1. *Let X and Y be random variables, and f and g be absolutely continuous functions such that* (2.7) *holds. Then*

$$\begin{aligned}
\mathrm{Cov}[f(X),g(Y)] &= \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]\right) \mathrm{d}f(x)\mathrm{d}g(y) && (3.2)\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]\right) f'(x)g'(y)\,\mathrm{d}x\mathrm{d}y, && (3.3)
\end{aligned}$$

where f' and g' are the derivatives of f and g which exist almost everywhere. The integrand $\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]$ in (3.2) *and* (3.3) *can be replaced by any of*

$$\begin{gathered}
\mathbb{P}[X \geq x, Y \geq y] - \mathbb{P}[X \geq x]\mathbb{P}[Y \geq y],\\
\mathbb{P}[X \geq x, Y > y] - \mathbb{P}[X \geq x]\mathbb{P}[Y > y],\\
\mathbb{P}[X > x, Y \geq y] - \mathbb{P}[X > x]\mathbb{P}[Y \geq y],
\end{gathered}$$

Proof. Write $f = f_1 - f_2$ and $g = g_1 - g_2$ for some non-decreasing absolutely continuous functions f_1, f_2, g_1, g_2. Due to the bilinearity of covariance,

$$\mathrm{Cov}[f(X), g(Y)] = \mathrm{Cov}[f_1(X), g_1(Y)] + \mathrm{Cov}[f_2(X), g_2(Y)] - \mathrm{Cov}[f_1(X), g_2(Y)] - \mathrm{Cov}[f_2(X), g_1(Y)].$$

Applying Corollary 2.7 (d) to the above four covariances, we have

$$\begin{aligned}
\mathrm{Cov}[f(X), g(Y)] &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y])\, \mathrm{d}f_1(x)\mathrm{d}g_1(y) \\
&\quad + \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y])\, \mathrm{d}f_2(x)\mathrm{d}g_2(y) \\
&\quad - \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y])\, \mathrm{d}f_1(x)\mathrm{d}g_2(y) \\
&\quad - \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y])\, \mathrm{d}f_2(x)\mathrm{d}g_1(y) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y])\, \mathrm{d}f(x)\mathrm{d}g(y).
\end{aligned}$$

This proves (3.2). By virtue of absolutely continuity, we may replace $\mathrm{d}f(x)$ and $\mathrm{d}g(y)$ by $f'(x)\,\mathrm{d}x$ and $g'(y)\,\mathrm{d}y$ respectively (see, e.g., Remark (v)(a) on page 285 of Stein and Shakarchi (2005)), arriving at (3.3).

To show the last part of the theorem, we note that the following bivariate functions

$$\begin{aligned}
(x, y) &\mapsto \mathbb{P}[X \geq x, Y \geq y] - \mathbb{P}[X \geq x]\mathbb{P}[Y \geq y] \\
(x, y) &\mapsto \mathbb{P}[X \geq x, Y > y] - \mathbb{P}[X \geq x]\mathbb{P}[Y > y] \\
(x, y) &\mapsto \mathbb{P}[X > x, Y \geq y] - \mathbb{P}[X > x]\mathbb{P}[Y \geq y] \\
(x, y) &\mapsto \mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]
\end{aligned}$$

are equal almost everywhere with respect to the Lebesgue measure on $\mathbb{R}^2$, and the induced measures of the absolutely continuous functions f_1, f_2, g_1, g_2 are also absolutely continuous with respect to the Lebesgue measure (see Theorem 39.12 of Aliprantis and Burkinshaw (1998)). Therefore, the term $\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]$ in (3.2) and (3.3) can be replaced by any of

$$\begin{aligned}
&\mathbb{P}[X \geq x, Y \geq y] - \mathbb{P}[X \geq x]\mathbb{P}[Y \geq y], \\
&\mathbb{P}[X \geq x, Y > y] - \mathbb{P}[X \geq x]\mathbb{P}[Y > y], \\
&\mathbb{P}[X > x, Y \geq y] - \mathbb{P}[X > x]\mathbb{P}[Y \geq y],
\end{aligned}$$

without altering the values of the integrals. □

Remark 3.2. (i) In the case where one of f and g is non-decreasing and the other one is absolutely continuous, the integral formula for $\mathrm{Cov}[f(X), g(Y)]$ consists of two terms. To see this, without loss of generality we assume that f is non-decreasing and g is absolutely continuous. Writing $f = f_r + f_l$ and $g = g_1 - g_2$ for some non-decreasing right-continuous function f_r, non-decreasing left-continuous function f_l and non-decreasing absolutely continuous functions

g_1 and g_2, we apply (3.1) and Corollary 2.7 (b) and (d) to obtain

$$\begin{aligned}
\mathrm{Cov}[f(X), g(Y)] &= \mathrm{Cov}[f_r(X), g_1(Y)] + \mathrm{Cov}[f_l(X), g_1(Y)] \\
&\quad - \mathrm{Cov}[f_r(X), g_2(Y)] - \mathrm{Cov}[f_l(X), g_2(Y)] \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X \geq x, Y > y] - \mathbb{P}[X \geq x]\mathbb{P}[Y > y]) \,\mathrm{d}f_r(x)\mathrm{d}g_1(y) \\
&\quad + \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]) \,\mathrm{d}f_l(x)\mathrm{d}g_1(y) \\
&\quad - \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X \geq x, Y > y] - \mathbb{P}[X \geq x]\mathbb{P}[Y > y]) \,\mathrm{d}f_r(x)\mathrm{d}g_2(y) \\
&\quad - \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]) \,\mathrm{d}f_l(x)\mathrm{d}g_2(y) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X \geq x, Y > y] - \mathbb{P}[X \geq x]\mathbb{P}[Y > y]) \,\mathrm{d}f_r(x)\mathrm{d}g(y) \\
&\quad + \int_{\mathbb{R}} \int_{\mathbb{R}} (\mathbb{P}[X > x, Y > y] - \mathbb{P}[X > x]\mathbb{P}[Y > y]) \,\mathrm{d}f_l(x)\mathrm{d}g(y).
\end{aligned}$$

Because g_1 and g_2 are absolutely continuous (in particular, left- and right-continuous), every "$Y > y$" in the preceding formula can also be replaced by "$Y \geq y$".

(ii) One can indeed extend Hoeffding's lemma to functions of bounded variation, as noted in Cuadras (2002), by decomposing such functions into differences of non-decreasing functions, to which Theorem 2.4 can be applied. The resulting formula for general random variables consists of a total of 16 non-recombining double integrals with different integrands and integrators. Due to the notational complexity of this ultimate formula and that there is no immediate need for such generality in the remainder of this paper and in most applications, we choose not to present it here.

4 A formula for Kendall's tau for general random variables

In the same spirit as Sections 2 and 3, this section provides an integral expression for Kendall's tau (Kendall (1938)), which is a nonparametric measure of concordance, for general random variables with possibly discontinuous distributions and a caveat when the discontinuity of distributions is not appropriately taken into account.

Definition 4.1. For a given random vector (X, Y), *Kendall's tau* is defined as the probability of concordance less the probability of discordance:

$$\tau[X, Y] := \mathbb{P}[(X - X')(Y - Y') > 0] - \mathbb{P}[(X - X')(Y - Y') < 0],$$

where (X', Y') is any i.i.d. copy of (X, Y).

In the literature, the treatment of Kendall's tau is almost exclusively restricted to random variables with continuous and strictly increasing distributions, for which the technicalities arising from changes of variables in integration can be sidestepped. In Theorem 4.2 below, we manage to overcome these technical challenges and generalize the conventional formula of Kendall's tau for continuous random variables to general, possibly discontinuous ones. The integrands of our integral formula parallel those of (2.8). An expression in a similar vein, stated in terms of distribution functions, can be found in Proposition 2.2 of Mesfioui and Tajar (2005).

Theorem 4.2. *For any random variables X and Y with joint distribution function F,*

$$\begin{aligned}\tau[X,Y] \;&=\; \int_{\mathbb{R}^2} \mathbb{P}[X \geq x, Y \geq y] \,\mathrm{d}F(x,y) + \int_{\mathbb{R}^2} \mathbb{P}[X \geq x, Y > y] \,\mathrm{d}F(x,y) \\ &\quad + \int_{\mathbb{R}^2} \mathbb{P}[X > x, Y \geq y] \,\mathrm{d}F(x,y) + \int_{\mathbb{R}^2} \mathbb{P}[X > x, Y > y] \,\mathrm{d}F(x,y) - 1. \qquad (4.1)\end{aligned}$$

Proof. We first write

$$\begin{aligned}\tau[X,Y] \;&=\; \mathbb{P}[(X-X')(Y-Y') > 0] - \mathbb{P}[(X-X')(Y-Y') < 0] \\ &=\; \mathbb{P}[X < X', Y < Y'] + \mathbb{P}[X > X', Y > Y'] - \mathbb{P}[X < X', Y > Y'] - \mathbb{P}[X > X', Y < Y'].\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{P}[X < X', Y > Y'] \;&=\; \mathbb{P}[X < X'] - \mathbb{P}[X < X', Y \leq Y'], \\ \mathbb{P}[X > X', Y < Y'] \;&=\; \mathbb{P}[Y < Y'] - \mathbb{P}[X \leq X', Y < Y'], \\ \text{and} \quad \mathbb{P}[X < X', Y < Y'] - \mathbb{P}[X < X'] - \mathbb{P}[Y < Y'] \;&=\; -\mathbb{P}[X < X' \text{ or } Y < Y'] \\ &=\; \mathbb{P}[X \geq X', Y \geq Y'] - 1.\end{aligned}$$

Thus

$$\begin{aligned}\tau[X,Y] \;&=\; \mathbb{P}[X < X', Y < Y'] + \mathbb{P}[X > X', Y > Y'] \\ &\quad - \big(\mathbb{P}[X < X'] - \mathbb{P}[X < X', Y \leq Y']\big) - \big(\mathbb{P}[Y < Y'] - \mathbb{P}[X \leq X', Y < Y']\big) \\ &=\; \mathbb{P}[X \geq X', Y \geq Y'] + \mathbb{P}[X \leq X', Y < Y'] + \mathbb{P}[X < X', Y \leq Y'] + \mathbb{P}[X > X', Y > Y'] - 1,\end{aligned}$$

which, together with the fact that (X,Y) and (X',Y') are i.i.d., leads to (4.1). □

Remark 4.3. (i) When the joint distribution/survival function of (X,Y) is continuous, we retrieve from (4.1) the usual formula (see, e.g., Equation (2.41) on page 55 of Joe (2015))

$$\tau[X,Y] = 4\int_{\mathbb{R}^2} \mathbb{P}[X > x, Y > y] \,\mathrm{d}F(x,y) - 1 = 4\mathbb{E}[\bar{F}(X,Y)] - 1, \qquad (4.2)$$

where $\bar{F}$ is the joint survival function of (X,Y).

(ii) As an example where (4.1) and (4.2) differ, consider the case when X and Y are i.i.d. Bernoulli random variables with a success probability of 0.5. If $\bar{F}_X$ and $\bar{F}_Y$ denote the survival functions of X and Y, then

$$\bar{F}_X(x) = \bar{F}_Y(x) = \begin{cases} 1, & \text{if } x < 0, \\ 0.5, & \text{if } 0 \leq x < 1, \\ 0, & \text{if } 1 \leq x. \end{cases}$$

It follows from (4.1) and the independence between X and Y that

$$\begin{aligned}\tau[X,Y] \;&=\; \mathbb{P}[X \geq X']\mathbb{P}[Y \geq Y'] + \mathbb{P}[X \geq X']\mathbb{P}[Y > Y'] \\ &\quad + \mathbb{P}[X > X']\mathbb{P}[Y \geq Y'] + \mathbb{P}[X > X']\mathbb{P}[Y > Y'] - 1 \\ &=\; \mathbb{P}[X \geq X']\left(\mathbb{P}[Y \geq Y'] + \mathbb{P}[Y > Y']\right) + \mathbb{P}[X > X']\left(\mathbb{P}[Y \geq Y'] + \mathbb{P}[Y > Y']\right) - 1 \\ &=\; \mathbb{P}[X' \geq X]\left(\mathbb{P}[Y' \geq Y] + \mathbb{P}[Y > Y']\right) + \mathbb{P}[X > X']\left(\mathbb{P}[Y' \geq Y] + \mathbb{P}[Y > Y']\right) - 1 \\ &=\; \mathbb{P}[X' \geq X] + \mathbb{P}[X > X'] - 1 \\ &=\; 0,\end{aligned}$$

whereas (4.2) gives

$$\tau[X,Y] = 4\mathbb{E}[\bar{F}_X(X)\bar{F}_Y(Y)] - 1 = 4\mathbb{E}[\bar{F}_X(X)]\mathbb{E}[\bar{F}_Y(Y)] - 1 = 4[0.5(0.5) + 0(0.5)]^2 - 1 = -0.75.$$

The discrepancy between the two values is a consequence of the discontinuity of the Bernoulli distribution.

Acknowledgments

The author would like to thank two anonymous reviewers and his colleague, Qihe Tang, at the University of Iowa for their exceptionally careful reading and valuable comments which have definitely improved the quality of the manuscript. Support from a start-up fund provided by the College of Liberal Arts and Sciences, The University of Iowa, and a Centers of Actuarial Excellence (CAE) Research Grant (2013–2016) from the Society of Actuaries is also gratefully acknowledged.

References

Aliprantis, C.D., Burkinshaw, O., 1998. Principles of Real Analysis. Academic Press. Third edition.

Cuadras, C.M., 2002. On the covariance between functions. Journal of Multivariate Analysis 81, 19–27.

Cuadras, C.M., 2015. Contributions to the diagonal expansion of a bivariate copula with continuous extensions. Journal of Multivariate Analysis 139, 28–44.

Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R., 2005. Actuarial Theory for Dependent Risks: Measures, Orders and Models. John Wiley & Sons, Inc.

Dhaene, J., Kukush, A., Linders, D., Tang, Q., 2012. Remarks on quantiles and distortion risk measures. European Actuarial Journal 2, 319–328.

Hoeffding, W., 1940. Masstabinvariante Korrelationstheorie. Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin 5, 181–233.

Joe, H., 2015. Dependence Modeling with Copulas. CRC Press, Florida.

Kallenberg, O., 2002. Foundations of Modern Probability. Springer-Verlag, New York. Second edition.

Kendall, M.G., 1938. A new measure of rank correlation. Biometrika 30, 81–93.

Leadbetter, R., Cambanis, S., Pipiras, V., 2014. A Basic Course in Measure and Probability. Cambridge University Press, New York.

Mardia, K.V., 1967. Some contributions to contingency-type bivariate distributions. Biometrika 54, 235–249.

Mesfioui, M., Tajar, A., 2005. On the properties of some nonparametric concordance measures in the discrete case. Journal of Nonparametric Statistics 17, 541–554.

Sen, P.K., 1994. The impact of Wassily Hoeffding's research on nonparametrics, in: The Collected Works of Wassily Hoeffding. Springer.

Stein, E.M., Shakarchi, R., 2005. Real Analysis: Measure Theory, Integration, and Hilbert Spaces. Princeton University Press, Princeton, New Jersey.