Meta Distributions—Part 2: Properties and Interpretations

Martin Haenggi, Fellow, IEEE

(Invited Paper)

Abstract—In the companion letter [1], we have defined and exemplified meta distributions (MDs) as a natural extension of the concepts of the mean and distribution of a random variable. Here we provide an in-depth discussion of the properties and interpretations of MDs. It includes original results on the calculation of MDs in the monotone case and two applications to simple Poisson wireless networks models.

Index Terms—Meta distributions, wireless networks, stochastic geometry, point processes, interference.

I. INTRODUCTION

The examples in Part 1 [1] have revealed a certain structure in the calculation of MDs. In this second part, we first derive a simple formula that applies to MDs of random variables of the form \( Z = f(X, Y) \) for (strictly) monotone functions \( f \) and complementary cumulative distributions (ccdfs) \( F_X \) and \( F_Y \). The formula leads to a simple proof that \( \bar{F}_Z(Y|X)(z, \cdot) \) and \( \bar{F}_Z(X|Y)(z, \cdot) \) are mutual inverses. Its application to two simple types of Poisson networks shows interesting connections to more complete models. Next, we offer several interpretations of the MD, including a dual interpretation that is based on a switch of the two parameters of the MD. Lastly we present an extension to higher-order MDs.

II. A FORMULA FOR THE GENERAL MONOTONE CASE

In this section we present a simple formula for the MD that applies to the monotone case.

Let \( f: (\mathbb{R}^+)^2 \to \mathbb{R}^+ \) be strictly monotone in both arguments and denote the inverse w.r.t. the first argument by \( f_1^{-1}(\cdot, y) \) and w.r.t. the second argument by \( f_2^{-1}(x, \cdot) \), i.e., \( f_1^{-1}(\cdot, y) \circ f(\cdot, y) = \mathbb{I} \) and \( f_2^{-1}(\cdot, x) \circ f(\cdot, x) = \mathbb{I} \), where \( \mathbb{I} \) is the identity operator on \( \mathbb{R}^+ \).

For instance, for \( z = f(x, y) = x/(x+y) \in [0,1] \), the inverses are

\[
\begin{align*}
  f_1^{-1}(z, y) &= \frac{zy}{1-z}, \\
  f_2^{-1}(z, x) &= \frac{x(1-z)}{z}, \quad z \in (0,1).
\end{align*}
\]

\( z = 1 \) implies \( y = 0 \) and \( x \) is arbitrary, while \( z = 0 \) implies \( x = 0 \) and \( y \) is arbitrary.

Theorem 1 If \( f(\cdot, y) \) and \( f(\cdot, \cdot) \) are both increasing, then

\[
\begin{align*}
  \bar{F}_Z(Y|X)(z, x) &= F_Y \left( f_2^{-1}(z, \bar{F}_X^{-1}(x)) \right) \\
  \bar{F}_Z(X|Y)(z, x) &= F_X \left( f_1^{-1}(z, \bar{F}_Y^{-1}(x)) \right). 
\end{align*}
\]

If \( f(\cdot, y) \) is increasing and \( f(x, \cdot) \) is decreasing,

\[
\begin{align*}
  \bar{F}_Z(Y|X)(z, x) &= F_Y \left( f_2^{-1}(z, \bar{F}_X^{-1}(x)) \right) \\
  \bar{F}_Z(X|Y)(z, x) &= F_X \left( f_1^{-1}(z, \bar{F}_Y^{-1}(x)) \right). 
\end{align*}
\]

Proof: See Appendix A.

So in Case 1, the two MDs are a composition of the three monotone functions \( F_Y \circ f_2^{-1}(z, \cdot) \circ \bar{F}_X^{-1} \) and \( F_X \circ f_1^{-1}(z, \cdot) \circ \bar{F}_Y^{-1} \), respectively. This also holds in the other two cases, with cdfs replaced by ccdfs or vice versa.

From the theorem follows a sufficient condition for the two MDs \( \bar{F}_Z(Y|X) \) and \( \bar{F}_Z(X|Y) \) to be inverses in the second argument.

Corollary 2 If \( f \) is monotone in both arguments, the two MDs \( \bar{F}_Z(Y|X)(z, \cdot) \) and \( \bar{F}_Z(X|Y)(z, \cdot) \) are mutual inverses, i.e., for all \( z \geq 0 \),

\[
\bar{F}_Z(Y|X)(z, y) = x \iff \bar{F}_Z(X|Y)(z, y) = x.
\]

Proof: By inspection and using the fact that for invertible functions \( g_1, g_2, \) and \( g_3 \), \( (g_1 \circ g_2 \circ g_3)^{-1} = g_3^{-1} \circ g_2^{-1} \circ g_1^{-1} \), we observe that the result holds in all three cases of increasing and decreasing \( f \).

III. APPLICATION TO SIMPLE WIRELESS NETWORK MODELS

Here we derive the SIR MD for two Poisson network models using the nearest-interferer-only approximation.

A. Poisson Cellular Downlink Networks

Corollary 3 Consider a downlink cellular network where BSs form a stationary PPP \( \Phi \subset \mathbb{R}^2 \) with nearest-BS association and Rayleigh fading. If only the nearest interfering BSs is considered, the SIR MD at an arbitrary location is

\[
\bar{F}_{\text{SIR}|x}(z, x) = \min \left\{ 1, \left( \frac{1-x}{x^\alpha} \right)^{\frac{1}{\alpha}} \right\}.
\]

Proof: The SIR is

\[
\text{SIR} = \frac{h_1}{h_2} \left( \frac{r_1}{r_2} \right)^{-\alpha},
\]
where \( h_k \) are the fading random variables, \( r_1 \) and \( r_2 \) are the distances to the serving and interfering BS, respectively. Letting \( X = h_1/h_2 \) and \( Y = (r_2/r_1)^\alpha \), \( X \) captures the fading, while \( Y \) captures the network geometry, and \( \text{SIR} = f(X, Y) = XY \). We have

\[
\bar{F}_X(x) = \frac{1}{1 + x}; \quad \bar{F}_Y(y) = \min\{1, y^{-\delta}\}.
\]

The ccdf of \( Y \) follows from [2, Lemma 1]. Hence \( \bar{F}_X^{-1}(x) = 1/x - 1 \) and \( f^{-1}(z, y) = z/y \). Now we have all ingredients to apply Theorem 1 and obtain

\[
\bar{F}_{[\text{SIR}|Y]}(z, x) = \bar{F}_Y \left( \bar{F}_X^{-1}(z, \bar{F}_X^{-1}(x)) \right)
= \min \left\{ 1, \left( \frac{z}{1/x - 1} \right)^{-\delta} \right\}.
\]

Remarks:

- Since the PPP only affects the SIR via \( Y \), \( \bar{F}_{[\text{SIR}|Y]} \), \( \bar{F}_{[\text{SIR}|Y]} \), and \( \bar{F}_{[\text{SIR}|Y]} \) are all equivalent.
- The MD in (7) provides an upper bound on the MD when all interferers are considered. This bound is not particularly tight for standard values of \( \delta \), \( z \), \( x \), but it gets tight as \( \delta \to 0 \) and provides the correct asymptotics \( \Theta(z^{-\delta}) \), \( z \to \infty \), and \( \Theta((1/x - 1)^\beta) \), \( x \to 1 \). For both \( z \to \infty \) and \( x \to 1 \), the pre-constant is 1 in (7), while it is since \( \delta = \sin(\pi\delta)/(\pi\delta) \) in the full-interference case [2, Cor. 5].
- Integration of (7) over \( x \) yields the SIR ccdf

\[
\bar{F}_{\text{SIR}}(z) = 2F_1(1, \delta; 1 + \delta, -z),
\]

where \( 2F_1 \) is the Gauss hypergeometric function. For \( \delta = 1/2 \), the SIR ccdf has the particularly simple form

\[
\bar{F}_{\text{SIR}}(z) = \arctan(\sqrt{z}/\sqrt{\bar{z}}).
\]

For comparison, in the full-interference case, the SIR ccdf is (3)

\[
\bar{F}_{\text{SIR}}(z) = \left( 2F_1(1, -\delta; 1 - \delta, -z) \right)^{-1}.
\]

- For \( x < 1/(1 + z) \), all users achieve reliability \( x \) for an SIR threshold \( z \) due to the ccdf of \( X \) and since multiplying with \( Y \) makes the SIR larger. Interestingly, the complementary region \( x > 1/(1 + z) \), where the MD is less than 1, corresponds to the separable region defined in [2, Thm. 1] and given in [2, Cor. 4] in the full-interference case. That theorem asserts that for \( (z, x) \) in the separable region, the MD in the full-interference case can be expressed as

\[
\bar{F}_{[\text{SIR}|X]}(z, x) = g(x) z^{-\delta}
\]

for any independent fading, where the function \( g(x) \) depends on the fading statistics and the path loss exponent. Hence we can conclude that (7) yields the upper bound \( g(x) < 1/(1 - 1)^{\delta} \).

Inverting the MD yields

\[
\bar{F}_{[\text{SIR}|X]}(z, x) = \frac{1}{1 + zx^{1/\delta}}, \quad x \in [0, 1), \quad \bar{F}_{[\text{SIR}|X]}(z, 1) = 0.
\]

This is the quantile function of (7), quantifying the performance of user percentiles. For instance, the 5% user's reliability\(^1\) for \( z = 1 \) and \( \alpha = 4 \) is \( 1/(1 + 0.95^2) \approx 0.53 \), while the 20% user's reliability is \( 1/(1 + 0.8^2) \approx 0.61 \).

Without fading, \( \bar{F}_X(x) \to 1(x < 1) \) and \( \bar{F}_X^{-1}(x) \) does not exist. In this case, \( U = 1((r_2/r_1)^\alpha > z) \in \{0, 1\} \), and the MD does not offer any extra information over just \( \bar{F}_Y \) since \( \text{SIR} = Y \) and \( \bar{F}_{Z|Y} \equiv \bar{F}_{Z|Z} = 1(Z > \cdot) \), where \( \equiv \) denotes an identity. It degenerates to \( \bar{F}_{[\text{SIR}|Y]}(z, x) = \min\{1, z^{-\delta}\} \).

For Nakagami-\(m \) fading, \( X \) has the F-distribution

\[
F_X(x) = I_x/(x+1)(m, m),
\]

where \( I_q(m, m) \) is the incomplete regularized beta function. The MD follows as

\[
\bar{F}_{[\text{SIR}|Y]}(z, x) = \min\left\{ 1, \left( \frac{q(x)}{z(1 - q(x))} \right)^{\delta} \right\},
\]

where \( q(x) = I_x^{-1}(m, m) \). For \( z > 1 \), \( \bar{F}_{[\text{SIR}|Y]}(z, 1)^2 = z^{-\delta} \) for all \( m > 0 \) since \( I_x^{-1}(m, m) = 1/2 \) irrespective of \( m \).

For \( m = 1/2 \), the F-distribution specializes to the beta prime distribution with parameters \( 1/2 \) and \( 1/2 \), given by \( F_X(x) = 2\arctan(x^{-1/2})/\pi \), leading to the closed-form expression

\[
\bar{F}_{[\text{SIR}|Y]}(z, x) = \min\{1, (\{z\tan(\pi x/2)^2\}^{-\delta} \}.
\]

Fig. 1 shows cross-sections of the MDs \( \bar{F}_{[\text{SIR}|Y]}(3, x) \) for different \( m \). It is apparent that for \( x = 1/2 \), the value for all \( m \) is \( 3^{-2/3} \approx 0.48 \).

B. Poisson Bipolar Networks

Next we consider a Poisson bipolar network where active transmitters form a PPP \( \Phi \subset \mathbb{R}^2 \) of intensity \( \lambda \) and have a dedicated receiver at distance 1. For general distances \( r \), the threshold \( z \) can simply be replaced by \( rz^\alpha \).

Corollary 4 Consider the Poisson bipolar model with link distance 1 and Rayleigh fading. If only the nearest interferer is considered, the SIR MD at the typical receiver is

\[
\bar{F}_{[\text{SIR}|Y]}(z, x) = \exp \left( -\lambda \pi \left( \frac{r x}{1-x} \right)^{\delta} \right). \tag{10}
\]

\(^1\)This is the reliability that 95% of the users achieve but 5% do not.
Replacing b derive results on the asymptotic behavior of MDs, in particular opens the door to the application of Tauberian theorems to in the high-reliability regime x.

\[ U \]

i.e. both full-interference case, whose curve is obtained by applying the Gil-Pelaez theorem to the imaginary moments given in [3, Thm. 1] by setting MD.

**Proof:** Let h be the fading coefficient from the desired transmitter and h1 from the nearest interferer, at distance r1 from the receiver. We apply Theorem 1 with \( X = h/h_1 \), \( Y = r_1^\alpha \), and SIR = XY. X is distributed as in the cellular case, while Y is Weibull distributed with ccdf \( F_Y(y) = e^{-\lambda y^\alpha} \). Combining the inverses of \( F_X \), \( f(x,y) = xy \), and using \( F_Y \) yields the result.

In contrast to the simple expression of the SIR MD, the SIR ccdf cannot be obtained in closed form.

In the case where all interferers are taken into account, the SIR ccdf is

\[ \bar{F}_{\text{SIR}}(z) = e^{-\lambda x z^\delta / \text{sinc} \delta}, \]

which follows as a special case of [3, Thm. 1] by setting \( b = p = 1 \). Hence \( \bar{F}_{\text{SIR}}(z, (1 + \text{sinc} \delta)^{-1}) = \bar{F}_{\text{SIR}}(z) \) and both \( \log \bar{F}_{\text{SIR}}(z) \) and \( \log \bar{F}_{\text{SIR}}(z, x) \) are proportional to \( z^\delta \).

Fig. 2 compares the MD in (10) with the MD of the SIR ccdf is

\[ S(z, x) = \sup_{p \in (0,1], \lambda > 0} p \lambda \bar{F}_{\text{SIR}(p, \lambda)}(z, x), \]

where \( \lambda \) is the density of \( \Phi \) and \( p \) is the fraction of active transmitters. The spatial outage capacity is the maximum density of links that a network can accommodate given an outage constraint.

**D. The Dual Interpretation**

Let \( U_z = \bar{F}_{\text{Z|Y}}(z) \), to make explicit that \( U \) is a function of \( z \). The MD \( \bar{F}_{\text{Z|Y}}(U) \) can be expressed as

\[ \bar{F}_{U_z}(x) = \mathbb{P}(\bar{F}_{\text{Z|Y}}(z) > x) = \mathbb{P}(z < \bar{F}_{\text{Z|Y}}^{-1}(x)) = \bar{F}_{U_z}(z), \]

where \( U_z = \bar{F}_{\text{Z|Y}}^{-1}(x) \) is the random value of \( z \) that achieves a reliability \( x \) for a given \( Y \). U and U are dual in the sense 2If there are multiple point processes involved, such as in a cellular network where base stations and users form point processes, the processes need to be jointly ergodic. Otherwise it is possible to construct examples where spatial averages do not correspond to ensembles averages, e.g., when base stations and users form square lattice of the same intensity, in which case all users have the same conditional SIR distribution (given the point processes).

3For instance the “user experienced data rate” in the specifications of the International Telecommunication Union (ITU) refers to the data rate of the 5% user.
Channels are subject to Rayleigh fading and path loss with exponent 2. The black number (at the top right) of each user is \( u \in \mathbb{P}(S_n > 1 \mid \Phi) \), and the blue number (at the bottom right) of each user is the value of \( z_u \) such that \( \mathbb{P}(S_n > z_u \mid \Phi) = 0.75 \).

that \( \mathbb{P}(U_z > x) = \mathbb{P}(U_z > x) \). This dual form of the MD was first reported in [5] and applied to a rate control problem in Poisson bipolar networks.

Integrating the MD over \( z \) yields the mean \( \mathbb{E}U_z \) (and further integration over \( x \) yields \( \mathbb{E}Z \)).

Let us re-visit the cellular uplink example in [1, Subs. II.A], where \( X = h + Y = R^2 \) are exponential with mean 1 and \( 1/(\lambda \pi) \), respectively, and \( Z = X/Y \). In this case, \( U_z = F_{Z|Y}(x) = -\log(x)/Y = -\log(x)/R^2 \). Fig. 3 shows the value of \( z \) for each user so that its link achieves a reliability of 0.75. The distribution of \( U_{3/4} = -\log(3/4)/R^2 \) follows as

\[
\mathbb{P}(U_{3/4} > z) = \mathbb{P}(U_z > 3/4) = F_{Z|Y}(z; 3/4) = 1 - \left(\frac{3}{4}\right) \lambda \pi / z.
\]

Hence the dual interpretation yields the distribution of the individual (per-user) threshold given a target reliability.

In the bipolar network in Subs. III-B, we have

\[
\mathbb{E}U_z = \frac{1}{1+z/Y}
\]

and thus \( U_z = Y(1/x - 1) \), with \( \mathbb{E}U_z(z) \) given in (10).

For instance, for a target reliability \( x = 10/11 \), \( \mathbb{P}(U_{10/11} > z) = e^{-10\lambda \pi z^2} \). The mean (for general \( x \)) is

\[
\mathbb{E}U_z = \frac{1 - x}{\Gamma(1+1/\delta)} \frac{\lambda \pi / z}{1+1/\delta}.
\]

Similarly, there exists \( V_x \) as the counterpart to \( V_z \), i.e., \( V_x = F_{Z|X}^{-1}(x) \) and \( \mathbb{P}(V_x > x) = \mathbb{P}(V_x > z) \).

V. Higher-order Meta Distributions

The concept of the MD is not restricted to two classes of randomness. For \( m \in \mathbb{N} \cup \{\infty\} \), let \( (X_j)_{j \in [m]} \) be a vector of random variables or random elements, and let \( \{P_1, \ldots, P_{n+1}\} \) be a partition of \([m]\). Then \( X_i \equiv (X_j)_{j \in [m]} \), with \( i \in [n + 1] \), splits the random vector in \( n + 1 \) classes.

**Definition 1 (n-th order meta distribution)** For \( n \in \mathbb{N}_0 \), let \( k \) denote the k-th permutation of \([n+1]\) in lexicographic order and denote the permuted numbers by \( k(1), k(2), \ldots, k(n+1) \). Let \( Z = f(X_1, \ldots, X_{n+1}) \). The n-th order meta distribution of \( Z \) with index \( k \), \( k \in ([n+1]!) \), is

\[
\bar{F}^{(n)}_{Z\mid k}(x) = \mathbb{P}(\mathbb{E}_{X_{k(n+1)}}(1 \cdots 1 \mathbb{E}_{X_{k(2)}}(1 \cdots 1 \mathbb{E}_{X_{k(1)}}(1 \mathbb{E}_{X_{k(1)}}(Z > z) > x_1) > x_2 \cdots ) > x_{n-1}) > x_n).
\]

The outermost expectation could equivalently be written as \( \mathbb{E}_{X_{k(n+1)}} \). Lower-order MDs can be obtained by integration over \( x_n, x_{n-1}, \ldots \).

In this general notation, the MDs considered in the previous sections and in Part I [1] are the first-order MDs \( \bar{F}^{(1)}_{Z\mid Y}(z, x) \equiv \mathbb{E}(Z)_{[1]}(z; x) \) and \( \bar{F}^{(1)}_{Z\mid X}(z, x) \equiv \mathbb{E}(Z)_{[1]}(z; x) \), and the ccdf of \( Z \) is the 0-th order MD

\[
\bar{F}_Z(z) = \bar{F}^{(0)}_{Z\mid X}(z).
\]

VI. Concluding Remarks

Meta distributions provide more refined information about the structure of random variables, much in the same way as distributions give more insight than averages. In the companion paper [1], we have presented several simple examples of the form \( Z = f(X, Y) \). This case has wide applicability since it is straightforward to extend \( f(X, Y) \) to \( g(f(g_1(X), g_2(Y))) \) by standard transformations of ccdfs. In fact, for suitable \( g_1, g_2 \), \( f(X, Y) = X + Y \) covers all functions of two variables. The examples have also shown that the MD is not necessarily more complicated than the ccdfs—in fact, the opposite is true in several cases.

In this second part we have established that under monotonicity conditions, the MDs \( \bar{F}^{(1)}_{Z\mid Y}(z, \cdot) \) and \( \bar{F}^{(1)}_{Z\mid X}(z, \cdot) \) are mutual inverses and thus also mutual quantile functions. It will be interesting to explore if this inverse property holds for larger classes of functions \( f \) and ccdfs of \( X \) and \( Y \).

Two simple stochastic geometry-based applications demonstrated how MDs achieve a time-scale separation between fading and spatial randomness and how closed-form expressions can be obtained that serve as bounds to more complete models and accurately reflect their asymptotic behavior. The averaging over the fading random variables while preserving the spatial randomness permits the definition of a robust notion of coverage in a cellular network as the event \( \{\text{SIR} > z\} \) (whose probability is the MD \( \bar{F}^{(1)}_{Z\mid Y}(z, x) \)). “Robust” here means that the coverage event does not depend on the instantaneous realization of the fading random variables as the frequently considered event \( \{\text{SIR} > z\} \) does (whose probability is the ccdf \( \bar{F}^{(1)}_{Z\mid X}(z) \)).

We call the branch of stochastic geometry that analyzes meta distributions deep stochastic geometry, to emphasize that the metrics derived provide a deeper analysis of the performance relative to the standard spatial averages (means and ccdfs).
The duality property of the MD is not solely of theoretical interest but has important applications in rate control when a fixed target link reliability is given. It gives the distribution of the threshold parameter $z$ over the network when the parameter $x$ in the MD is set to a fixed value.

Higher-order MDs constitute a natural extension of the concept. While conceptually straightforward, it remains to be seen whether they have any impactful applications, for instance when the time scales of small-scale fading, shadowing, and node locations are separated, or in spatiotemporal models where additional time scales come into play due to dynamic and scheduling. Other future work may include the study of the effect of noise in the models, finding the necessary condition for the inversion property in Cor. 2, identifying the limits of the beta approximation of MDs, and exploring whether considering dependence in the classes $X_i$ would be leading to new insights.

APPENDIX

A. Proof of Theorem 1

We first prove a lemma relating the inverses $f_1^{-1}$ and $f_2^{-1}$.

**Lemma 1** The two inverses $f_1^{-1}(z, \cdot)$ and $f_2^{-1}(z, \cdot)$ are mutual inverses, i.e., $f_1^{-1}(z, \cdot) \circ f_2^{-1}(z, \cdot) = \mathbb{I}$.

**Proof:** We need to show that, for all $z \geq 0$,

$$f_1^{-1}(z, f_2^{-1}(z, x)) = x = f_2^{-1}(z, f_1^{-1}(z, x)).$$

The definitions of $f_1^{-1}$ and $f_2^{-1}$ imply that for all $y \geq 0$,

$$f(f_1^{-1}(z, y), y) = z,$$

and for all $x \geq 0$,

$$f(x, f_2^{-1}(z, x)) = z.$$ (14)

In particular, (14) holds for $y = f_2^{-1}(z, x)$, i.e.,

$$f(f_1^{-1}(z, f_2^{-1}(z, x)), f_2^{-1}(z, x)) = z,$$

which, from (15), implies that $x = f_1^{-1}(z, f_2^{-1}(z, x))$.

Similarly, (15) holds for $x = f_1^{-1}(z, y)$ and thus

$$f(f_1^{-1}(z, y), f_2^{-1}(z, f_1^{-1}(z, y))) = z.$$

Next, we proceed to the proof of the theorem.

**Proof:** We consider the three cases of increasing and decreasing $f$ separately.

**Case 1.** If $f(\cdot, y)$ and $f(x, \cdot)$ are both increasing, we have

$$P(f(X, Y) > z \mid Y) = P(X > f_1^{-1}(z, Y) \mid Y) = F_X(f_1^{-1}(z, Y))$$

$$P(f(X, Y) > z \mid X) = P(Y > f_2^{-1}(z, X) \mid Y) = F_Y(f_2^{-1}(z, X))$$

and thus, taking into account that the inversion of cdfs results in a change in the direction of the inequality,

$$F_{[Z|Y]}(z, x) = P(F_X(f_1^{-1}(z, Y)) > x)$$

$$= P(f_1^{-1}(z, Y) < F_X^{-1}(x))$$

$$(a)\ P(Y > f_2^{-1}(z, F_X^{-1}(x)))$$

$$= F_Y(f_2^{-1}(z, F_X^{-1}(x))),$$

where step (a) uses Lemma 1 for the inversion, which changes the direction of the inequality.

**Case 2.** If $f(\cdot, y)$ is increasing but $f(x, \cdot)$ is decreasing, we have

$$P(f(X, Y) > z \mid Y) = P(X > f_1^{-1}(z, Y) \mid Y) = F_X(f_1^{-1}(z, Y))$$

$$P(f(X, Y) > z \mid X) = P(Y < f_2^{-1}(z, X) \mid Y) = F_Y(f_2^{-1}(z, X))$$

and

$$P(F_X(f_1^{-1}(z, Y)) > x) = P(F_Y(f_2^{-1}(z, Y)) < F_Y^{-1}(x))$$

$$(a)\ P(Y < f_2^{-1}(z, F_Y^{-1}(x)))$$

$$= F_Y(f_2^{-1}(z, F_Y^{-1}(x))),$$

where step (a) uses Lemma 1 for the inversion, which changes the direction of the inequality.

**Case 3.** If $f(\cdot, y)$ and $f(x, \cdot)$ are both decreasing, we have

$$P(f(X, Y) > z \mid Y) = P(X < f_1^{-1}(z, Y) \mid Y) = F_X(f_1^{-1}(z, Y))$$

$$P(f(X, Y) > z \mid X) = P(Y < f_2^{-1}(z, X) \mid Y) = F_Y(f_2^{-1}(z, X))$$

and

$$P(F_X(f_1^{-1}(z, Y)) > x) = P(f_1^{-1}(z, Y) > F_X^{-1}(x))$$

$$(a)\ P(Y < f_2^{-1}(z, F_X^{-1}(x)))$$

$$= F_Y(f_2^{-1}(z, F_X^{-1}(x))),$$

where step (a) uses Lemma 1 for the inversion, which changes the direction of the inequality.

REFERENCES


