An easily verifiable dispersion order for discrete distributions

Andreas Eberl[1], Bernhard Klar[1], Alfonso Suárez-Llorens[2]

[1]Karlsruhe Institute of Technology [2]Universidad de Cádiz

Published in *REVSTAT – Statistical Journal* (2025)

Stochastic models in biomathematics and applications
20-21 January, 2026.
University of Salerno, Fisciano (SA), Italy

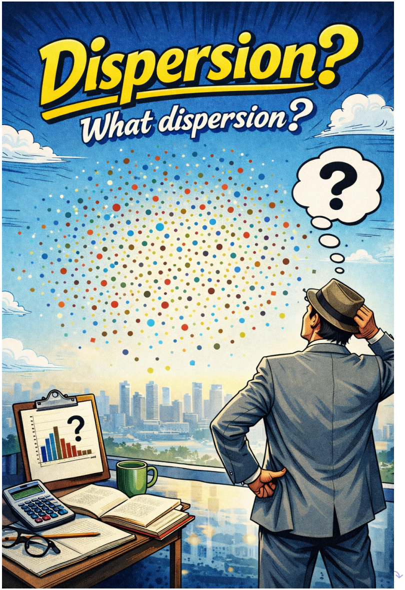
PID2024-157464NB-I00, funded by MCIU (AEI), and European Union

Dispersion? What dispersion?

Dispersion is a fundamental concept in Statistics. It plays a crucial role...

However, what do we mean when we speak about *dispersion*?

- Given a r.v X and a functional $\nu(X)$, under what conditions can the mapping $\nu(X)$ be considered a valid measure of dispersion?
- Given two r.v.s X and Y, under what conditions can we say that X is less dispersed than Y in the sense of a stochastic comparison?



Question 1: What is a measure of dispersion?

What properties should any dispersion measure $\nu(X)$ satisfy?

- **Law invariance:** if $X =_{st} Y$, then $\nu(X) = \nu(Y)$.
- **Non-negativity:** $\nu(X) \geq 0$, with equality if X is degenerate.
- **Translation invariance:** $\nu(X + cte) = \nu(X)$.
- **Absolute homogeneity:** $\nu(\lambda X)) = |\lambda| \, \nu(X)$.

Some classical examples:

- **The standard deviation:** $SD(X) = \sqrt{E(X-\mu)^2}$.
- **The Gini mean difference:** $GMD(X) = E \mid X_1 - X_2 \mid$, where X_1 and X_2 are independent copies of X.
- **The mean absolute deviation:** $MAD(X) = E \mid X - \mu \mid$.

Bibliography.

Bickel and Lehmann [1976, 1979].

Question 2: How to compare the dispersion of X and Y?

Stochastic orderings provide a detailed way to compare r.v.s that goes beyond single numerical measures... We say that, $X \leq_* Y$.

- **Comparisons in variability**: *dispersive order*, $\leq_{\text{disp}}$, *convex order*, $\leq_{\text{cx}}$, *spread order*, $\leq_{\text{spread}}$, *dilatation order*, $\leq_{\text{dil}}$, etc.
- **Consistency:** We expect $X \leq_{\textit{variability}} Y \Rightarrow \nu(X) \leq \nu(Y)$.
- The **dispersive order** plays a central role, as it is the strongest variability ordering and dispersion measures are usually assumed to be consistent with it.

Dispersive order. We say that $X \leq_{\text{disp}} Y$ if the differences between corresponding quantiles of Y dominate those of X, i.e.,

$$F_X^{-1}(q) - F_X^{-1}(p) \leq F_Y^{-1}(q) - F_Y^{-1}(p), \quad \forall\, 0 < p < q < 1.$$

Biblography.

Shaked and Shanthikumar [2006], Müller and Stoyan [2002], Bickel and Lehmann [1976, 1979], Oja [1981], Lewis and Thompson [1981].

Dispersion orderings: the discrete case

A problematic point. Houston, we have a problem!

Dispersion measures and **variability orderings** usually work well for **continuous distributions** and have a clear interpretation. However, they become too restrictive in the **discrete case**.

- **Dispersive order:** A necessary condition for $X \leq_{\rm disp} Y$ is

$$\mathrm{Im}(F_X(\cdot)) \subseteq \mathrm{Im}(F_Y(\cdot)),$$

  (see Müller and Stoyan [2002] and Eberl and Klar [2025]).
- While **innocuous for continuous distributions**, this condition is **highly restrictive in the discrete case**.
- It excludes most **lattice** and **empirical distributions**.

The difficulty stems from **CDF discontinuities** and **irregular quantile spacing**, which male difficult quantile-based comparisons.

Why $\leq_{\text{disp}}$ is too restrictive in the discrete case

Too restrictive.

For discrete distributions, the necessary condition $\text{Im}(F) \subseteq \text{Im}(G)$ is often violated, even for very natural and comparable models.

- **Discrete uniform distributions**:

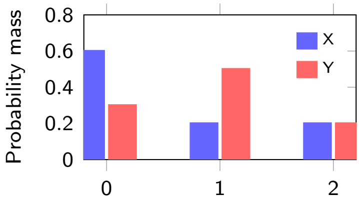
$$X \sim U\{1:4\} \text{ and } Y \sim U\{1:5\},$$

$$\text{Im}(F_X) = \left\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\} \not\subset \text{Im}(F_Y) = \left\{\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1\right\}.$$

- **Bernoulli distributions** with different success probabilities.
- **Poisson distributions** with different mean value.
- **Binomial distributions** with different parameters.
- **Geometric distributions** with different success probabilities.
- **Empirical distributions**.

Classical dispersion measures

Not only classical **variability orderings** may fail in the discrete setting, but also some widely used **dispersion measures**.

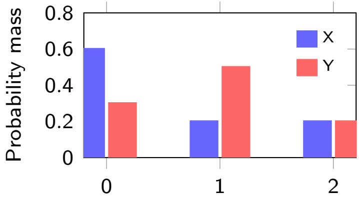


	SD	MAD	GMD
X	0.80	0.72	0.80
Y	0.65	0.54	0.74

- X should be considered *more dispersed* than Y?

Classical dispersion measures

Not only classical **variability orderings** may fail in the discrete setting, but also some widely used **dispersion measures**.

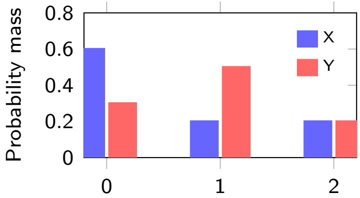


	SD	MAD	GMD
X	0.80	0.72	0.80
Y	0.65	0.54	0.74

- X should be considered *more dispersed* than Y?
- But, it seems that X is actually more *predictable*!

Classical dispersion measures

Not only classical **variability orderings** may fail in the discrete setting, but also some widely used **dispersion measures**.



	SD	MAD	GMD
X	0.80	0.72	0.80
Y	0.65	0.54	0.74

- X should be considered *more dispersed* than Y?
- But, it seems that X is actually more *predictable*!
- What about the **Shannon entropy**, $H(p) = -\sum p_i \log_2(p_i)$.

$$H(X) = 1.371 \quad \text{and} \quad H(Y) = 1.485.$$

The lower entropy of X indicates less uncertainty...

Discrete Dispersive Orders: Motivation & Approach

Why do we need new approaches in the discrete case?

Classical dispersion orders are too rigid for discrete distributions and fail to capture the essence of the probability mass function (pmf).

Key ideas for discrete-tailored orders.

The main challenge is bridging continuous smooth CDFs and discrete step-function CDFs, requiring flexible definitions of discrete dispersion

Practical benefits.

Easily verifiable on real-world count data, useful for modeling and detecting **overdispersion**, and provides a bridge between theory and discrete data.

We consider a classical concentration function

Definition. A classical function.

Let X be a real-valued random variable with distribution function F. The Lévy concentration function associated with X is the mapping $Q_X : \mathbb{R}^+ \longrightarrow [0,1]$ where

$$Q_X(\varepsilon) = \sup_{x_0 \in \mathbb{R}} \Pr\{X \in [x_0, x_0 + \varepsilon]\}.$$

Notes.

- Lévy [1937], " Théorie de l'addition des variables aléatoires".
- It is well-defined, right-continuous and non-decreasing in ε, where tends to 1 when ε goes to infinity.
- It was used by Fernández-Ponce and Suárez-Llorens [2003] for comparing continuous random variables.
- It has received little attention in stochastic order theory.

Our Weak Dispersive Order for Discrete Variables

Definition. Based on the concentration function.

Let X and Y be two discrete r.v.s. Then X is said to be *less weakly dispersive* than Y, denoted by $X \leq_{\text{wd}} Y$ if

$$\begin{aligned} Q_X(\varepsilon) &\geq Q_Y(\varepsilon), \forall \varepsilon > 0, \\ \sup_{x_0} \Pr\{X \in [x_0, x_0 + \varepsilon]\} &\geq \sup_{y_0} \Pr\{Y \in [y_0, y_0 + \varepsilon]\}, \forall \varepsilon > 0. \end{aligned}$$

Notes.

- $\leq_{\text{disp}}$ fixes the accumulated probability in the interquantile interval, $q - p$, and compares the interval length.
- $\leq_{\text{wd}}$ fixes the interval length, ε, and compares the probability.
- The supremum may not be achieved.

Properties of the Weak Dispersive Order

Some elementary properties.

- $X \leq_{\mathrm{wd}} Y$ formally requires infinitely many comparisons. In practice only a countable -or even finite- number is needed.
- If there exists $\varepsilon > 0$ that is smaller than any distance between points in the supports of X and Y, then

$$X \leq_{\mathrm{wd}} Y \Rightarrow \max_i \{p_i\} \geq \max_i \{q_i\}.$$

- If X and Y have bounded supports, then

$$X \leq_{\mathrm{wd}} Y \Rightarrow \max_i \{x_i\} - \min_i \{x_i\} \leq \max_i \{y_i\} - \min_i \{y_i\}.$$

Properties of the Weak Dispersive Order

We recall the definition of a contraction function.

Let $\phi : \mathbb{R} \mapsto \mathbb{R}$ be a real function. We will say that ϕ is a contraction if $|\phi(y) - \phi(x)| \leq |y - x|$, for all $x, y \in \mathbb{R}$.

Properties of the Weak Dispersive Order

We recall the definition of a contraction function.

Let $\phi : \mathbb{R} \mapsto \mathbb{R}$ be a real function. We will say that ϕ is a contraction if $|\phi(y) - \phi(x)| \leq |y - x|$, for all $x, y \in \mathbb{R}$.

Property 1.

Let X and Y be two random variables such that $X =_{st} \phi(Y)$, where ϕ is a monotone contraction function. Then $X \leq_{\text{wd}} Y$.

Properties of the Weak Dispersive Order

We recall the definition of a contraction function.

Let $\phi : \mathbb{R} \mapsto \mathbb{R}$ be a real function. We will say that ϕ is a contraction if $|\phi(y) - \phi(x)| \leq |y - x|$, for all $x, y \in \mathbb{R}$.

Property 1.

Let X and Y be two random variables such that $X =_{st} \phi(Y)$, where ϕ is a monotone contraction function. Then $X \leq_{\text{wd}} Y$.

Consequences of Property 1.

- $X \leq_{\text{disp}} Y \Rightarrow X \leq_{\text{wd}} Y$, as expected.
- $X \leq_{\text{wd}} (\geq_{wd})\, aX + b$, $\forall |a| > (<)\, 1$, $b \in \mathbb{R}$.
- $X =_{wd} -X + b =_{wd} X + b$, $\forall b \in \mathbb{R}$.
- $X =_{\text{wd}} -X$. This is not generally true for $\leq_{\text{disp}}$ or $\leq_{\text{cx}}$.

Properties of the Weak Dispersive Order

Property 2, Lévy [1937].

Let X and Y be two independent random variables. Then,

$$Q_X(\varepsilon) \geq Q_{X+Y}(\varepsilon), \text{ for all } \varepsilon > 0.$$

Consequently, $X \leq_{wd} X + Y$ holds.

Properties of the Weak Dispersive Order

Property 2, Lévy [1937].

Let X and Y be two independent random variables. Then,

$$Q_X(\varepsilon) \geq Q_{X+Y}(\varepsilon), \text{ for all } \varepsilon > 0.$$

Consequently, $X \leq_{wd} X + Y$ holds.

Consequences of Property 2.

1. Convolution with an independent random variable tends to increase dispersion.
2. This property does not generally hold for other dispersive ordering.
3. This property allows many possible comparisons.

Properties of the Weak Dispersive Order

We recall the definition of the usual stochastic order.

Let X and Y be two r.v.s. We say that X is smaller that Y in the usual stochastic order, denoted by $X \leq_{\text{st}} Y$, if

$$Prob(X \leq a) \geq Prob(Y \leq a), \text{ for all } a \in \mathbb{R}.$$

Properties of the Weak Dispersive Order

We recall the definition of the usual stochastic order.

Let X and Y be two r.v.s. We say that X is smaller that Y in the usual stochastic order, denoted by $X \leq_{\text{st}} Y$, if

$$Prob(X \leq a) \geq Prob(Y \leq a), \text{ for all } a \in \mathbb{R}.$$

Property 3.

Let X and Y be two r.v.s on $\mathbb{N}$ with decreasing pmfs. Then,

$$X \leq_{\text{st}} Y \Rightarrow X \leq_{\text{wd}} Y.$$

Properties of the Weak Dispersive Order

We recall the definition of the usual stochastic order.

Let X and Y be two r.v.s. We say that X is smaller that Y in the usual stochastic order, denoted by $X \leq_{\text{st}} Y$, if

$$Prob(X \leq a) \geq Prob(Y \leq a), \text{ for all } a \in \mathbb{R}.$$

Property 3.

Let X and Y be two r.v.s on $\mathbb{N}$ with decreasing pmfs. Then,

$$X \leq_{\text{st}} Y \Rightarrow X \leq_{\text{wd}} Y.$$

Consequences of Property 3.

1. There are many discrete distributions with a decreasing pmf.
2. This property also allows many possible comparisons.

Properties of the Weak Dispersive Order.

We recall the defintion of randomness, Hickey [1983]

Let X and Y be discrete r.v.s and let $p = \{p_i\}$ and $q = \{q_i\}$ be the pmfs defined on the union of their supports.

$$X \leq_{rand} Y \text{ if } \sum_{i=1}^{k} p_{(i)} \geq \sum_{i=1}^{k} q_{(i)}, \quad k = 1, \ldots, n,$$

where $p_{(i)}$ and $q_{(i)}$ denote the i-th largest values of p and q.

Properties of the Weak Dispersive Order.

We recall the defintion of randomness, Hickey [1983]

Let X and Y be discrete r.v.s and let $p = \{p_i\}$ and $q = \{q_i\}$ be the pmfs defined on the union of their supports.

$$X \leq_{rand} Y \text{ if } \sum_{i=1}^{k} p_{(i)} \geq \sum_{i=1}^{k} q_{(i)}, \quad k = 1, \ldots, n,$$

where $p_{(i)}$ and $q_{(i)}$ denote the i-th largest values of p and q.

We recall the definition of unimodality.

Let $p = \{p_i\}$ on $\mathbb{Z}$. Then, it is unimodal if $\exists M \in \mathbb{Z}$ such that

$$\ldots p_{M-2} \leq p_{M-1} \leq p_M \geq p_{M+1} \geq p_{M+1} \geq \ldots$$

Properties of the Weak Dispersive Order.

We recall the defintion of randomness, Hickey [1983]

Let X and Y be discrete r.v.s and let $p = \{p_i\}$ and $q = \{q_i\}$ be the pmfs defined on the union of their supports.

$$X \leq_{rand} Y \text{ if } \sum_{i=1}^{k} p_{(i)} \geq \sum_{i=1}^{k} q_{(i)}, \quad k = 1, \ldots, n,$$

where $p_{(i)}$ and $q_{(i)}$ denote the i-th largest values of p and q.

We recall the definition of unimodality.

Let $p = \{p_i\}$ on $\mathbb{Z}$. Then, it is unimodal if $\exists M \in \mathbb{Z}$ such that

$$\ldots p_{M-2} \leq p_{M-1} \leq p_M \geq p_{M+1} \geq p_{M+1} \geq \ldots$$

Property 4.

Let $p = \{p_i\}$ and $q = \{p_i\}$ be unimodal distributions on $\mathbb{N}_0$.

$$X \leq_{\text{wd}} Y \Leftrightarrow X \leq_{rand} Y \stackrel{\text{Entropies}}{\Longrightarrow} H(X) \leq H(Y).$$

Examples of ordered classical families

Example (A straightforward computation.)

- Uniform r.v.s: $U\{1:m\} \leq_{\text{wd}} U\{1:n\}$, for $m \leq n$
- Bernouilli r.v.s: $Be(p) \leq_{\text{wd}} Be(0.5)$, $\forall p \in [0,1]$.

$$Be(p_1) \leq_{\text{wd}} Be(p_2) \Leftrightarrow \max\{p_1, 1-p_1\} \geq \max\{p_2, 1-p_2\}.$$

Example (As a direct consequence of Properties 2 or 4.)

- Poisson r.v.s: $P(\lambda_1) \leq_{\text{wd}} P(\lambda_2) \iff \lambda_1 \leq \lambda_2$.
- Binomial r.v.s: $B(m,p) \leq_{\text{wd}} B(n,p) \iff m \leq n$.
- Negative binomial r.v.s: $NB(r,p) \leq_{\text{wd}} NB(s,p)Y \iff r \leq s$.
- Hermite r.v.s: $H(a,b) \leq_{\text{wd}} H(c,d) \iff a \leq c$ and $b \leq d$.

Example (As a direct consequence of Property 3.)

- Geometric r.v.s: $Ge(p_1) \leq_{\text{wd}} Ge(p_2) \iff p_1 \geq p_2$.
- Logarithmic r.v.s: $Log(p_1) \leq_{\text{wd}} Log(p_2) \iff p_1 \leq p_2$.

Relationships with classical dispersion measures

A key inequality.

Let X be a r.v, it is easy to check that

$$E[|X-a|^r] \geq r \int_0^\infty t^{r-1}(1-Q_X(2t))dt, \quad \forall a \in \mathbb{R}, \forall r \geq 1.$$

Therefore,

$$E[|X-a_r|^r] \geq \frac{r}{2^r} \int_0^\infty \varepsilon^{r-1}(1-Q_X(\varepsilon))\, d\varepsilon, \quad \forall r \geq 1,$$

where $a_r = \arg\min_a E[|X-a|^r]$, $r \geq 1$.

Note.

The funtion $Q_X(\varepsilon)$ leads to a lower bound for several classical measures of variability (those that quantify deviation from a central point).

A new dispersion measure

Definition.

Any functional $\nu_r(X)$ defined as

$$\nu_r(X) = \frac{\sqrt[r]{r}}{2}\left(\int_0^\infty \varepsilon^{r-1}(1 - Q_X(\varepsilon))d\varepsilon\right)^{1/r},$$

can also be used as a measure of the variability of X.

Note. Interpretation.

It is essentially a weighted sum of tail probabilities outside intervals of length ε (small values indicate high concentration).

Note. A dispersion measure in the sense of Bickel and Lehmann.

Of course, $\nu_r(X)$ satisfies the classical properties of law invariance, non-negativity, translation invariance, and absolute homogeneity, as well as being consistent with the orderings $\leq_{\text{wd}}$ and $\leq_{\text{disp}}$.

A new dispersion measure

Note. Sometimes $\nu_r(X)$ is not difficult to compute.

For a random variable X with values in $\mathbb{N}_0$, we obtain

$$\nu_r(X) = \frac{1}{2}\left(\sum_{k=0}^{\infty}((k+1)^r - k^r)(1 - Q_X(k))\right)^{1/r}.$$

and, in particular, $\nu_1(X) = 1/2\sum_{k=0}^{\infty}(1 - Q_X(k))$.

Example

- $\nu_1(Be(p)) = \frac{1}{2}\min\{p, 1-p\}, \quad \nu_2(B(p)) = \frac{1}{2}\sqrt{\min\{p, 1-p\}}$
- $\nu_1(Ge(p)) = \frac{1-p}{2p}, \quad \nu_2(Ge(p)) = \frac{\sqrt{(1-p)(2-p)}}{2p}$
- $\nu_1(U\{1:m\}) = \frac{m-1}{4}, \quad \nu_2(U\{1:m\}) = \frac{1}{2}\sqrt{\frac{(M-1)(2M^2+5M-6)}{3M}}$

Other properties and concepts

Relationship with the dispersive order given by Eberl and Klar [2025]

Let X and Y with distribution functions as defined as in Eberl and Klar [2025, Def. 3.1]. Then, $X \leq_{\text{disp}}^{\wedge\text{-disc}} Y$ implies $X \leq_{\text{wd}} Y$. This new order provides a canonical tool for comparing dispersion in discrete r.v.s.

Counterexamples

All implications of $\leq_{\text{wd}}$ for $\leq_{\text{disp}}$, $\leq_{\text{rand}}$, and $\leq_{\text{disp}}^{\wedge\text{-disc}}$ are strict.

Robustness with respect to outliers in the empirical distribution

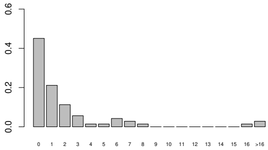
Let $X_1, \ldots, X_n$ be an i.i.d. sample from X. The plug-in estimator of $\nu_r(X)$ is non-robust, but analogous robust measures are easily defined.

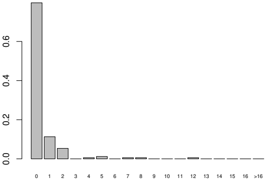
What is the equivalence class of a distribution with respect to $=_{wd}$?

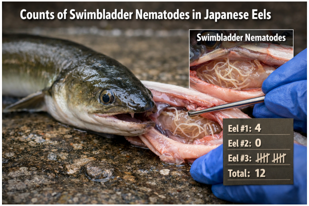
Let X and Y be two r.v.s having support $\{1, \ldots, 5\}$ and pmfs: $p = (0.1, 0.4, 0.05, 0.3, 0.15)$ and $q = (0.4, 0.1, 0.25, 0.15, 0.1)$. Then $X =_{wd} Y$, but they are not location shifts of each other.

Example. Data from Münderle et al. [2006]

Swimbladder nematode counts in Japanese eels (number of worm parasites).







Eels	Size	$\bar{x}$	$\widehat{\mathrm{SD}}$	$\widehat{\mathrm{CV}}$	$\widehat{\mathrm{MAD}}$	$\widehat{\mathrm{GMD}}$	$\widehat{\mathrm{IQR}}$	$\widehat{\nu}_1$	$\widehat{\nu}_2$
Aquacultured (Left)	71	3.32	9.39	2.83	4.29	5.48	2	1.65	4.95
Wild (Right)	168	0.46	1.43	3.11	0.74	0.83	0	0.23	0.75

Checking orderings for empirical cumulative distribution funtions (ECDF)

The ecdfs satisfies $F_{\text{wild},m} \leq_{\text{wd}} F_{\text{aqua},n}$ (no other dispersive orders hold) and $F_{\text{wild},m} \leq_{\text{st}} F_{\text{aqua},n}$. As expected, wild eels show both lower parasite counts and lower dispersion, reflecting lower and more uniform parasite exposure, whereas cultured eels face higher transmission under crowded conditions.

Future Research Directions

- Characterize equivalence classes under weak dispersive order, especially non-unimodal distributions.
- Develop multivariate discrete extensions using metric neighborhoods (e.g., Euclidean balls). Apply to spatial count data.
- Investigate concentration as a lower bound for variability measures; explore links with convex and dilation orders.
- Order new families of discrete distributions.
- Explore its effectiveness as a decision tree splitting criterion compared to Gini and entropy.

La fine, vi aspetto alla gelateria Nettuno

References I

P. J. Bickel and E. L. Lehmann. Descriptive statistics for nonparametric models. iii. dispersion. *The Annals of Statistics*, 4:1139–1158, 1976.

P. J. Bickel and E. L. Lehmann. Descriptive statistics for nonparametric models IV. Spread. in: J. Jurečková, Ed. In *Contributions to Statistics, Jaroslav Hájek Memorial Volume*. 33-40, Academia, Prague, 1979.

A. Eberl and B. Klar. Defining dispersion: A fundamental order for univariate discrete distributions. *ALEA - Latin American Journal of Probability and Mathematical Statistics*, 22(2):859–888, 2025.

J. M. Fernández-Ponce and A. Suárez-Llorens. An aging concept based on majorization. *Probability in the Engineering and Informational Sciences*, 17 (1):107–117, 2003.

R. J. Hickey. Majorisation, randomness and some discrete distributions. *Journal of Applied Probability*, 20(4):897–902, 1983.

References II

T. Lewis and J. W. Thompson. Dispersive distributions, and the connection between dispersivity and strong unimodality. *Journal of Applied Probability*, 18(1):76–90, 1981.

P. Lévy. *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris, 1937.

A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics. Wiley, 2002.

M. Münderle, H. Taraschewski, B. Klar, C. W. Chang, J. C. Shiao, K. N. Shen, J. T. He, S. H. Lin, and W. N. Tzeng. Occurrence of Anguillicola crassus (Nematoda: Dracunculoidea) in Japanese eels Anguilla japonica from a river and an aquaculture unit in southwest Taiwan. *Diseases of Aquatic Organisms*, 71:101–108, 2006.

H. Oja. On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics*, 8:154–168, 1981.

M. Shaked and J. G. Shanthikumar. *Stochastic Orders*. Springer Series in Statistics. Springer, 2006.