

On Conditional Distributions of Nearest Neighbors

E. KAUFMANN AND R.-D. REISS

FB 6, Uni-Gh Siegen, Hölderlinstrasse 3, D-5900 Siegen, Germany

Communicated by the Editors

Let X_1, \dots, X_n be i.i.d. S -valued random variables and let g be a real-valued function on S . We give an explicit representation of the conditional distribution of the empirical point process based on X_1, \dots, X_n given the $(k+1)$ th smallest order statistic of the r.v.'s $g(X_1), \dots, g(X_n)$. The extension to conditioning on several of the order statistics of $g(X_1), \dots, g(X_n)$ is indicated. The result for point processes enables us to deduce the conditional distribution of the k smallest g -order statistics taken in the order of their magnitude as well as in the order of their outcome. The latter r.v.'s are conditionally independent. © 1992 Academic Press, Inc.

1. INTRODUCTION

Let X be an S -valued r.v. and let X_1, \dots, X_n be independent replicas of X . Moreover, let $g: S \rightarrow \mathbb{R}$ be a measurable function such that $g(X)$ has a continuous d.f. Putting

$$x \leqslant_g y \quad \text{if} \quad g(x) \leqslant g(y)$$

for $x, y \in S$ one defines the g -ordering on S . The g -order statistics $X_{i:n}$ corresponding to X_1, \dots, X_n have the property

$$X_{1:n} \leqslant_g \dots \leqslant_g X_{n:n}$$

(see, e.g., Reiss [10, Section 2.1]). If $S = \mathbb{R}$ and g is the identity, this definition leads to the usual notion of order statistics on the real line.

Rearranging the k smallest g -order statistics $X_{1:n}, \dots, X_{k:n}$ in the original order of their outcome one obtains the r.v.'s $Z_{1,n}, \dots, Z_{k,n}$. Denote by

$$g_{1:n}(X) \leqslant \dots \leqslant g_{n:n}(X)$$

Received December 18, 1990; revised November 21, 1991.

AMS 1980 classification numbers: 60G55, 62G30.

Key words and phrases: conditional distribution, nearest neighbor, empirical point processes, g -order statistics.

the order statistics of $g(X_1), \dots, g(X_n)$ on the real line. It is clear that $g_{k:n}(X) = g(X_{k:n})$.

We will study the conditional distributions

$$P((X_{1:n}, \dots, X_{k:n}) \in \cdot \mid g_{k+1:n}(X) = r), \quad (1)$$

$$P((Z_{1:n}, \dots, Z_{k:n}) \in \cdot \mid g_{k+1:n}(X) = r), \quad (2)$$

and

$$P(N_n \in \cdot \mid g_{k+1:n}(X) = r), \quad (3)$$

where N_n is the empirical process $\sum_{i=1}^n \varepsilon_{X_i}$ with ε_x denoting the Dirac measure at the point x ; that is $\varepsilon_x(B) = 1_B(x)$ for $B \in \mathcal{B}$.

Our result may be regarded as a multivariate extension of a well-known result for univariate order statistics and exceedances (see Reiss [10, Theorem 1.8.1 and Corollary 1.8.4 for a rigorous proof). Such extensions have frequently been applied in the statistical literature. We mention applications in nonparametric density estimation (Moore and Yackel [9] and Hall [6]) with $g(x) = \|x - x_0\|$, where $\|\cdot\|$ denotes the Euclidean distance, in non-parametric regression analysis (Mack [7], Devroye [4], and Falk and Reiss [5]) with $g(x, y) = \|x - x_0\|$, and an application in the context of trimmed means (Maller [8]). The existing proofs are very short and it seems to be difficult to remove their heuristic character (at least, in the general framework). One may add that the results in the literature are restricted to the Euclidean space, whereas the present one holds if the r.v.'s X_i are measurable with respect to a Borel σ -field \mathcal{B} on a Polish space S .

In Section 2, explicit representations of the conditional distributions (1) and (2) will be deduced from that in (3). Our basic result, namely a representation of (3) as the distribution of a sum of three independent point processes, will be given in Section 3. Using point processes we avoid technical difficulties that come up in the context of ordering.

2. CONDITIONAL DISTRIBUTIONS OF ORDERED R.V.'S

In the sequel, we assume that \mathcal{B} is the Borel- σ -field of a Polish space S and that $g(X)$ has a continuous d.f. Hence, in our context, conditional distributions will always exist. It is understood that we deal with certain specific versions of the conditional distributions or the corresponding factorizations.

Our first theorem concerns the k smallest g -order statistics $X_{1:n} \leq_g \dots \leq_g X_{k:n}$ and the r.v.'s $Z_{1:n}, \dots, Z_{k:n}$. Note that

$$(Z_{1:n}, \dots, Z_{k:n}) = (X_{i(1)}, \dots, X_{i(k)}) \quad (4)$$

whenever $1 \leq i(1) < \dots < i(k) \leq n$ and $\max(g(X_{i(1)}), \dots, g(X_{i(k)})) < g_{k+1:n}(X)$. This defines $(Z_{1:n}, \dots, Z_{k:n})$ with probability one because $g(X)$ has a continuous d.f.

THEOREM 1. *Assume $g(X)$ has a continuous d.f. and that*

$$P\{g(X) < r\} > 0.$$

Denote by $Y_{1:k}^r \leq_g \dots \leq_g Y_{k:k}^r$ the g -order statistics of k i.i.d. random variables Y_1^r, \dots, Y_k^r with common distribution $P(X \in \cdot | g(X) < r)$. Then, for $k = 1, \dots, n-1$,

$$P((X_{1:n}, \dots, X_{k:n}) \in \cdot | g_{k+1:n}(X) = r) = \mathcal{L}(Y_{1:k}^r, \dots, Y_{k:k}^r) \quad (5)$$

and

$$P((Z_{1:n}, \dots, Z_{k:n}) \in \cdot | g_{k+1:n}(X) = r) = \mathcal{L}(Y_1^r, \dots, Y_k^r). \quad (6)$$

Proof. From Corollary 1 in Section 3 we obtain

$$P(N_{k,n} \in \cdot | g_{k+1:n}(X) = r) = \mathcal{L}\left(\sum_{i=1}^k \varepsilon_{Y_i^r}\right), \quad (7)$$

where

$$N_{k,n} = \sum_{i=1}^k \varepsilon_{X_{i:n}}. \quad (8)$$

Denote by $x_{1:n} \leq_g \dots \leq_g x_{n:n}$ the g -ordered values of $x_1, \dots, x_n \in S$ whenever $g(x_i) \neq g(x_j)$, $i \neq j$. Let $T_n: S^n \rightarrow S^n$ be a measurable map such that

$$T_n(x_1, \dots, x_n) = (x_{1:n}, \dots, x_{n:n})$$

whenever $g(x_i) \neq g(x_j)$, $i \neq j$. We may write

$$(X_{1:n}, \dots, X_{n:n}) = T_n(X_1, \dots, X_n).$$

Check that \tilde{T}_n defined by

$$\tilde{T}_n\left(\sum_{i=1}^n \varepsilon_{x_i}\right) := T_n(x_1, \dots, x_n)$$

is a measurable map on the space of point measures having exactly n points. Obviously,

$$\tilde{T}_k(N_{k,n}) = (X_{1:n}, \dots, X_{k:n})$$

and, hence,

$$\begin{aligned}
 P((X_{1:n}, \dots, X_{k:n}) \in \cdot \mid g_{k+1:n}(X) = r) \\
 &= P(N_{k,n} \in \tilde{T}_k^{-1}(\cdot) \mid g_{k+1:n}(X) = r) \\
 &= \mathcal{L} \left(\tilde{T}_k \left(\sum_{i=1}^k \varepsilon_{Y_i^r} \right) \right) \\
 &= \mathcal{L}(Y_{1:k}^r, \dots, Y_{k:n}^r).
 \end{aligned}$$

The proof of (5) is complete.

Denote by \mathcal{S}_k the group of permutations τ on $\{1, \dots, k\}$. In a routine way we may prove that

$$P((Y_1^r, \dots, Y_k^r) \in A \mid (Y_{1:k}^r, \dots, Y_{k:n}^r)) = \frac{1}{k!} \sum_{r \in \mathcal{S}_k} 1_A(Y_{\tau(1):k}^r, \dots, Y_{\tau(k):k}^r)$$

and R_n and $(X_{1:n}, \dots, X_{n:n})$ are independent, where R_n denotes the rank statistic pertaining to $g(X_1), \dots, g(X_n)$. Hence, (6) may be verified by using the same arguments as in the proof of Corollary 1.8.4 in Reiss [10]. ■

We will briefly discuss two applications of Theorem 1. The first concerns nonparametric regression—that is, estimation of functionals of the conditional d.f.—based on the k nearest neighbors of a fixed point x_0 (as already indicated in the introduction). Given i.i.d. bivariate random vectors $Z_i = (X_i, Y_i)$, $i = 1, \dots, n$, denote by $Z_{i:n}$ the g -order statistics where $g(x, y) = \|x - x_0\|$. It is proved in [5] that the Y -values pertaining to $Z_{i:n} \leq_g \dots \leq_g Z_{k:n}$ may approximately be treated like i.i.d. random variables with common d.f. being equal to the conditional d.f. of Y_1 given $X_1 = x_0$. The decisive step towards that result is accomplished by applying Theorem 1.

Another example for g -ordering is provided by “concomitants of order statistics” that are also termed “induced order statistics” (see e.g. [2, 1, 3]). If g is the projection into the first component and, hence, bivariate random vectors $(X_1, Y_1), \dots, (X_n, Y_n)$ are ordered according to the magnitude of the first component we obtain

$$(X_{1:n}, Y_{[1:n]}) \leq_g \dots \leq_g (X_{n:n}, Y_{[n:n]}),$$

where $Y_{[i:n]}$ is the i th concomitant. Theorem 1 implies that, conditioned on $X_{k+1:n} = r$, the concomitants $(Y_{[1:n]}, \dots, Y_{[k:n]})$ are distributed like concomitants $(Y'_{[1:k]}, \dots, Y'_{[k:k]})$ of i.i.d. random vectors $(X'_1, Y'_1), \dots, (X'_k, Y'_k)$ with common distribution $P((X_1, Y_1) \in \cdot \mid X_1 < r)$.

3. CONDITIONAL DISTRIBUTIONS OF POINT PROCESSES

It is shown that the conditional distribution of $N_n = \sum_{i=1}^n \varepsilon_{X_i}$ given the order statistic $g_{k+1:n}(X) = r$ is equal to the distribution of the sum of three independent empirical processes having mass only on $g^{-1}((-\infty, r])$, $g^{-1}(\{r\})$ and, respectively, $g^{-1}([r, \infty))$.

THEOREM 2. *If $g(X)$ has a continuous d.f. then, for $k = 1, \dots, n-1$,*

$$P(N_n \in \cdot \mid g_{k+1:n}(X) = r) = \mathcal{L}(N_n^r), \quad (9)$$

where $N_n^r = \sum_{i=1}^n \varepsilon_{Y_i^r}$ is a point process based on the independent r.v.'s Y_1^r, \dots, Y_n^r that are distributed as

$$\begin{aligned} P(X \in \cdot \mid g(X) < r) & \quad 1 \leq i \leq k, P\{g(X) < r\} > 0, \\ \mathcal{L}(Y_i^r) = P(X \in \cdot \mid g(X) > r) & \quad \text{if } k+2 \leq i \leq n, P\{g(X) > r\} > 0, \\ P(X \in \cdot \mid g(X) = r) & \quad \text{otherwise.} \end{aligned} \quad (10)$$

According to Theorem 2,

$$P(N_n \in \cdot \mid g_{k+1:n}(X) = r) = \mathcal{L}\left(N_k^r + \varepsilon_{Y_{k+1}^r} + \sum_{i=k+2}^n \varepsilon_{Y_i^r}\right),$$

where N_k^r is the empirical process $\sum_{i=1}^k \varepsilon_{Y_i^r}$ with Y_i^r , $i = 1, \dots, k$, given in (10).

We also formulate an extension of Theorem 2, where the empirical point process N_n is conditioned on two or more order statistics.

THEOREM 3. *Suppose that $g(X)$ has a continuous d.f. For $1 \leq m \leq n$ and $1 \leq k_1 < \dots < k_m < k_{m+1} = n+1$ we have*

$$P(N_n \in \cdot \mid g_{k_1:n}(X) = r_1, \dots, g_{k_m:n}(X) = r_m) = \mathcal{L}(N_n^{r_1 \dots r_m}),$$

where $N_n^{r_1 \dots r_m} = \sum_{i=1}^n \varepsilon_{Y_i^{r_1 \dots r_m}}$ and $Y_i^{r_1 \dots r_m}$, $i = 1, \dots, n$, are independent and distributed as follows:

$$\begin{aligned} P(X \in \cdot \mid g(X) < r_1) & \quad 1 \leq i < k_1, P\{g(X) < r_1\} > 0, \\ P(X \in \cdot \mid g(X) = r_1) & \quad 1 \leq i < k_1, P\{g(X) < r_1\} = 0, \\ P(X \in \cdot \mid g(X) \in (r_j, r_{j+1})) & \quad \text{if } k_j < i < k_{j+1}, P\{g(X) \in (r_j, r_{j+1})\} > 0, \\ P(X \in \cdot \mid g(X) = r_j) & \quad k_j < i < k_{j+1}, P\{g(X) \in (r_j, r_{j+1})\} = 0, \\ P(X \in \cdot \mid g(X) = r_j) & \quad i = k_j, \end{aligned}$$

where $r_{m+1} := \infty$ by convention.

The proof of Theorem 3 runs the same way as that of Theorem 2, yet due to the technical nature of the problem, the representation would be more complicated. The point process

$$N_{k,n} = \sum_{i=1}^k \varepsilon_{X_{i:n}}$$

may be represented by the empirical process N_n truncated outside of $D(g_{k+1:n}(X))$, where $D(r) = \{x: g(x) < r\}$; that is,

$$N_{k,n} = N_n(\cdot \cap D(g_{k+1:n}(X))).$$

The conditional distribution of $N_{k,n}$ given $g_{k+1:n}(X) = r$ will be deduced from Theorem 2. Note that for r.v.'s Z, Y and a measurable map f the conditional distribution of $f(Z, Y)$ given $Z = z$ is given by

$$\int 1_{(\cdot, \cdot)}(f(z, y)) P(Y \in dy \mid Z = z). \quad (11)$$

COROLLARY 1. *We have*

$$P(N_{k,n} \in \cdot \mid g_{k+1:n}(X) = r) = \mathcal{L}(N_k^r),$$

where N_k^r is the point process in Theorem 2.

Proof. Theorem 2 and formula (11) will be applied to $Z = g_{k+1:n}(X)$, $Y = N_n$, and

$$f(r, \mu) = \mu(\cdot \cap D(r))$$

for $r \in \mathbb{R}$ and point measures μ . Note that

$$f(Z, Y) = N_{k,n}.$$

Therefore, for every measurable set M of point measures,

$$\begin{aligned} & P(N_{k,n} \in M \mid g_{k+1:n}(X) = r) \\ &= \int 1_M(\mu(\cdot \cap D(r))) P(N_n \in d\mu \mid g_{k+1:n}(X) = r) \\ &= \int 1_M(\mu) \mathcal{L}\left(\left(N_n^r + \varepsilon_{Y_{k+1}^r} + \sum_{i=k+2}^n \varepsilon_{Y_i^r}\right)(\cdot \cap D(r))\right)(d\mu) \\ &= \mathcal{L}(N_k^r)(M). \end{aligned}$$

The proof is complete. ■

In the sequel we use the abbreviations

$$g_n(X) := (g(X_1), \dots, g(X_n))$$

and

$$g_{:n}(X) := (g_{1:n}(X), \dots, g(X_{n:n})).$$

Proof of Theorem 2. Define by $\iota_n(x) = \sum_{i=1}^n \varepsilon_{x_i}$, $x \in S^n$, a map from S^n onto the set of point measures on S having n points. Moreover, $\iota_n Q$ is the probability measure induced by ι_n and Q ; that is $\iota_n Q = Q(\iota_n^{-1}(\cdot))$. Put $u = (u_1, \dots, u_n) \in \mathbb{R}^n$.

The assertion (9) of the theorem will be deduced from the following three auxiliary results:

$$P(N_n \in \cdot | g_{:n}(X) = u) = P(N_n \in \cdot | g_n(X) = u), \quad (12)$$

$$P(N_n \in \cdot | g_n(X) = u) = \iota_n \bigg(\prod_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i) \bigg) \quad (13)$$

for $\mathcal{L}(g_{:n}(X))$ —a.a. $u \in \mathbb{R}^n$, and

$$\int \left(\iota_n \bigg(\prod_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i) \bigg) \right) \mathcal{L}(g_{:n}(Y^r))(du) = \mathcal{L}(N_n^r) \quad (14)$$

for $\mathcal{L}(g_{k+1:n}(X))$ —a.a. $r \in \mathbb{R}$.

To prove (9) we also apply Theorem 1.8.1 in Reiss [10], where it is shown that

$$P(g_{:n}(X) \in \cdot | g_{k+1:n}(X) = r) = \mathcal{L}(g_{:n}(Y^r)). \quad (15)$$

For $B \in \mathbb{B}$, with $\pi_k(u) = u_k$, $u \in \mathbb{R}^n$, denoting the k th projection, we obtain from (12)–(15),

$$\begin{aligned} & \int_B P(N_n \in \cdot | g_{k+1:n}(X) = r) \mathcal{L}(g_{k+1:n}(X))(dr) \\ &= \mathcal{L}(N_n, g_{k+1:n}(X))(\cdot \times B) \\ &= \int_{\pi_k^{-1}(B)} P(N_n \in \cdot | g_{:n}(X) = u) \mathcal{L}(g_{:n}(X))(du) \\ &\stackrel{(12)}{=} \int_{\pi_k^{-1}(B)} P(N_n \in \cdot | g_n(X) = u) \mathcal{L}(g_{:n}(X))(du) \\ &\stackrel{(13)}{=} \int_{\pi_k^{-1}(B)} \left(\iota_n \bigg(\prod_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i) \bigg) \right) \mathcal{L}(g_{:n}(X))(du) \end{aligned}$$

$$\begin{aligned}
&= \int_B \left(\int \left(\iota_n \bigotimes_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i) \right) \right. \\
&\quad \left. P(g_{:n}(X) \in du | g_{k+1:n}(X) = r) \right) \mathcal{L}(g_{k+1:n}(X))(dr) \\
&\stackrel{(15)}{=} \int_B \left(\int \left(\iota_n \bigotimes_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i) \right) \right. \\
&\quad \left. \mathcal{L}(g_{:n}(Y^r))(du) \right) \mathcal{L}(g_{k+1:n}(X))(dr) \\
&\stackrel{(14)}{=} \int_B \mathcal{L}(N_n^r)(\cdot) \mathcal{L}(g_{k+1:n}(X))(dr).
\end{aligned}$$

This yields the desired assertion (9), namely, that

$$P(N_n \in \cdot | g_{k+1:n}(X) = r) = \mathcal{L}(N_n^r)$$

for $\mathcal{L}(g_{k+1:n}(X))$ —a.a. $r \in \mathbb{R}$.

It remains to verify (12)–(14). Let \mathcal{S}_n again denote the group of permutations. We will also write $\tau(u) = (u_{\tau(1)}, \dots, u_{\tau(n)})$ for $u \in \mathbb{R}^n$. To prove (12) note that for every $\tau \in \mathcal{S}_n$ we obtain $P(N_n \in \cdot | g_n(X) = u) = P(N_n \in \cdot | g_n(X) = \tau(u))$ for $\mathcal{L}(g_n(X))$ —a.a. $u \in \mathbb{R}^n$ and, as well, for $\mathcal{L}(g_{:n}(X))$ —a.a. $u \in \mathbb{R}^n$, since the latter measure is dominated by $\mathcal{L}(g_n(X))$.

Because

$$P(g_n(X) \in \cdot | g_{:n}(X) = u) = \frac{1}{n!} \sum_{\tau \in \mathcal{S}_n} \varepsilon_{\tau(u)}$$

the Chapman–Kolmogorov equation yields

$$\begin{aligned}
&P(N_n \in \cdot | g_{:n}(X) = u) \\
&= \int P(N_n \in \cdot | g_n(X) = v) P(g_n(X) \in dv | g_{:n}(X) = u) \\
&= \frac{1}{n!} \sum_{\tau \in \mathcal{S}_n} P(N_n \in \cdot | g_n(X) = \tau(u)) \\
&= P(N_n \in \cdot | g_n(X) = u)
\end{aligned}$$

for $\mathcal{L}(g_{:n}(X))$ —a.a. $u \in \mathbb{R}^n$. The proof of (12) is complete.

Next, using standard arguments and the independence of $(X_1, g(X_1))$, ..., $(X_n, g(X_n))$, it is easily seen that

$$\begin{aligned}
P(N_n \in \cdot | g_n(X) = u) &= \iota_n P((X_1, \dots, X_n) \in \cdot | g_n(X) = u) \\
&= \iota_n \bigotimes_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i)
\end{aligned}$$

$\mathcal{L}(g_{:n}(X))$ —a.s., and (13) is verified. The last equation is a slight generalization of Lemma 1 in Bhattacharya [1].

Finally, we verify (14). Note that $\mathcal{L}(Y^r)$ is a Markov kernel and thus also $\mathcal{L}(N_n^r) = \iota_n \mathcal{L}(Y_1^r, \dots, Y_n^r)$ and $\mathcal{L}(g_{:n}(Y^r))$. We may write for $\mathcal{L}(g_{k+1:n}(X))$ —a.a. $r \in \mathbb{R}$,

$$P(X_i \in \cdot | g(X_i) = u) = P(Y_i^r \in \cdot | g(Y_i^r) = u) \quad \text{for } \mathcal{L}(g(Y_i^r))\text{—a.a. } u \in \mathbb{R}^n$$

and, by using the same arguments as above, we deduce for $\mathcal{L}(g_{k+1:n}(X))$ —a.a. $r \in \mathbb{R}$ and for $\mathcal{L}(g_n(Y^r))$ —a.a. $u \in \mathbb{R}^n$,

$$\begin{aligned} \prod_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i) &= \prod_{i=1}^n P(Y_i^r \in \cdot | g(Y_i^r) = u_i) \\ &= P((Y_1^r, \dots, Y_n^r) \in \cdot | g_n(Y^r) = u). \end{aligned}$$

Since $\mathcal{L}(g_n(Y^r))$ dominates $\mathcal{L}(g_{:n}(Y^r))$ we also obtain

$$\begin{aligned} &\int \left(\iota_n \prod_{i=1}^n P(X_i \in \cdot | g(X_i) = u_i) \right) \mathcal{L}(g_{:n}(Y^r))(du) \\ &= \int (\iota_n P((Y_1^r, \dots, Y_n^r) \in \cdot | g_n(Y^r) = u)) \mathcal{L}(g_{:n}(Y^r))(du) \\ &= \int P(N_n^r \in \cdot | g_n(Y^r) = u) \mathcal{L}(g_{:n}(Y^r))(du) \\ &= \int P(N_n^r \in \cdot | g_{:n}(Y^r) = u) \mathcal{L}(g_{:n}(Y^r))(du) \\ &= \mathcal{L}(N_n^r) \end{aligned}$$

for $\mathcal{L}(g_{k+1:n}(X))$ —a.a. $r \in \mathbb{R}$. The proof is complete. ■

REFERENCES

- [1] BHATTACHARYA, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics. *Ann. Statist.* **2** 1034–1039.
- [2] DAVID, H. A. (1981). *Order Statistics*, 2nd ed. Wiley, New York.
- [3] DAVID, H. A., AND GALAMBOS, J. (1974). The asymptotic theory of concomitants of order statistics. *J. Appl. Probab.* **11** 762–770.
- [4] DEVROYE, L. (1982). Necessary and sufficient conditions for the pointwise convergence of nearest neighbor regression function estimates. *Z. Wahrsch. Verw. Gebiete* **61** 467–481.
- [5] FALK, M., AND REISS, R.-D. (submitted). Estimating conditional curves: Nearest neighbor approach.

- [6] HALL, P. (1983). On near neighbour estimates of a multivariate density. *J. Multivariate Anal.* **13** 24–39.
- [7] MACK, Y. P. (1981). Local properties of k -NN regression estimates. *SIAM J. Algebraic Discete Methods* **2** 311–323.
- [8] MALLER, R. A. (1988). Asymptotic normality of trimmed means in higher dimensions. *Ann. Probab.* **16** 1608–1622.
- [9] MOORE, D. S., AND YACKEL, J. W. (1977). Large sample properties of nearest neighbor density function estimates. In *Statistical Decision Theory and Related Topics* (S. S. Gupta and D. S. Moore, Eds.), pp. 269–279. Academic Press, New York.
- [10] REISS, R.-D. (1989). *Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics*. New York, Springer.