Applications of the delta method with the least concave majorant operator

Brendan K. Beare

Department of Economics, University of California – San Diego

Banff International Research Station, January 30, 2018

Based on work with Zheng Fang (Texas A&M) and Jong-Myun Moon (PIMCO).

Overview

- The asymptotic behavior of statistics constructed from a least concave majorant (LCM) can sometimes be studied using the delta method.
- The LCM operator is not Hadamard differentiable but satisfies a weaker notion of smoothness sufficient to apply the delta method.
- The delta method for the bootstrap does not apply (Dümbgen, 1993).
- Application #1: We study a test of the null hypothesis that the ratio of two PDFs is monotone.
- Application #2: We study the behavior of the antiderivative of the Grenander estimator and associated resampling methods.
- Application #3: We briefly consider the application of our results to isotonic regression.

The least concave majorant

Definition

The least concave majorant is the operator

$$\mathcal{M} : \ell^{\infty}(\mathbf{R}^{+}) \to \ell^{\infty}(\mathbf{R}^{+})$$

that maps each $\theta \in \ell^{\infty}(\mathbf{R}^{+})$ to the function

$$\mathcal{M}\theta(x) = \inf\{g(x) : g \in \ell^{\infty}(\mathbf{R}^{+}), g \text{ is concave, and } \theta \leq g\}, \quad x \in \mathbf{R}^{+}.$$

Note that $\theta \leq \mathcal{M}\theta \leq \sup_{x \in \mathbf{R}^{+}} \theta(x)$, so we may take $\ell^{\infty}(\mathbf{R}^{+})$ to be the codomain of $\mathcal{M}$.

The least concave majorant

Definition

The least concave majorant over a nonempty convex set $T \subseteq \mathbf{R}^+$ is the operator

$$\mathcal{M}_T : \ell^\infty(\mathbf{R}^+) \to \ell^\infty(T)$$

that maps each $\theta \in \ell^\infty(\mathbf{R}^+)$ to the function

$$\mathcal{M}_T\theta(x) = \inf\{g(x) : g \in \ell^\infty(T), g \text{ is concave, and } \theta \leq g \text{ on } T\}, \quad x \in T.$$

Note that $\theta \leq \mathcal{M}_T\theta \leq \sup_{x \in T} \theta(x)$, so we may take $\ell^\infty(T)$ to be the codomain of $\mathcal{M}_T$.

Hadamard directional differentiability

Definition

Let $\mathbf{D}$ and $\mathbf{E}$ be Banach spaces. A map $\phi : \mathbf{D} \to \mathbf{E}$ is said to be Hadamard directionally differentiable at $\theta \in \mathbf{D}$ tangentially to a set $\mathbf{D}_0 \subset \mathbf{D}$ if there is a map $\phi'_\theta : \mathbf{D}_0 \to \mathbf{E}$ such that

$$\left\| \frac{\phi(\theta + t_n h_n) - \phi(\theta)}{t_n} - \phi'_\theta(h) \right\|_{\mathbf{E}} \to 0$$

for all $h \in \mathbf{D}_0$ and all $h_1, h_2, \ldots \in \mathbf{D}$ and $t_1, t_2, \ldots \in \mathbf{R}^+$ such that $t_n \downarrow 0$ and $\|h_n - h\|_{\mathbf{D}} \to 0$.

- Concept originates with Shapiro (1990,1991) and also used by Dümbgen (1993).
- Distinct from Hadamard differentiability because the approximating map ϕ'_θ need not be linear.
- The approximating map ϕ'_θ is always continuous and positive homogeneous of degree one.
- Hadamard directional differentiability is sufficient to apply the delta method (Shapiro, 1991, 1992) but not the delta method for the bootstrap (Dümbgen, 1993). See also Fang and Santos (2016).

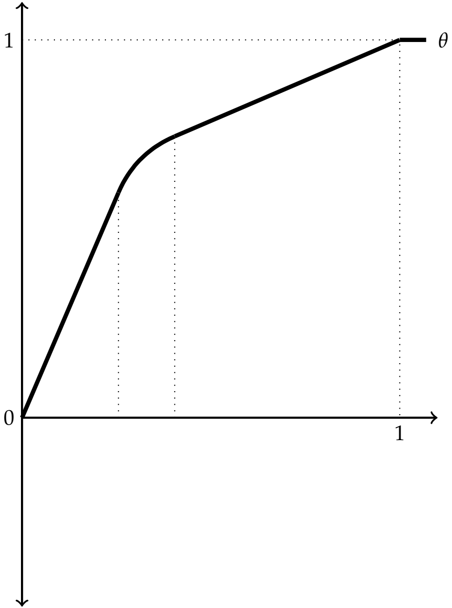
Directional derivative of the least concave majorant

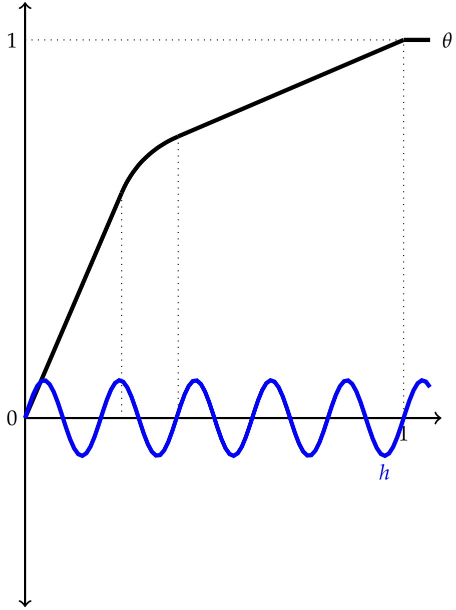
- Let $C_0(\mathbf{R}^+) \subset \ell^\infty(\mathbf{R}^+)$ be the continuous real functions on $\mathbf{R}^+$ vanishing at infinity, equipped with the uniform metric.

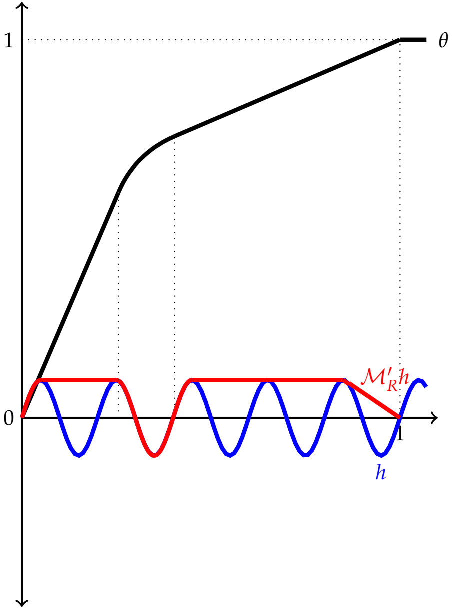
Theorem

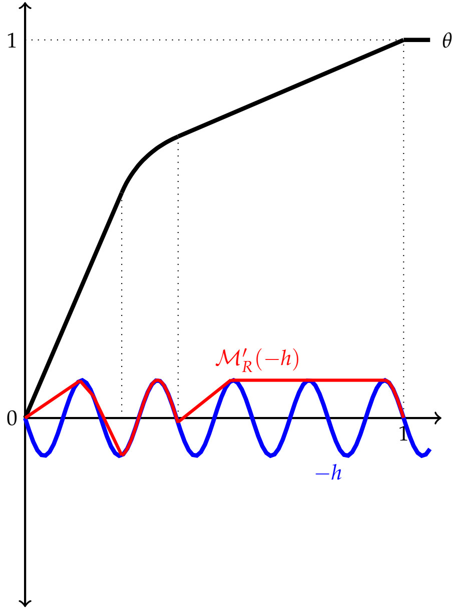
The LCM operator $\mathcal{M} : \ell^\infty(\mathbf{R}^+) \to \ell^\infty(\mathbf{R}^+)$ is Hadamard directionally differentiable at any concave $\theta \in \ell^\infty(\mathbf{R}^+)$ tangentially to $C_0(\mathbf{R}^+)$. Its directional derivative $\mathcal{M}'_\theta : C_0(\mathbf{R}^+) \to \ell^\infty(\mathbf{R}^+)$ is uniquely determined as follows: for any $h \in C_0(\mathbf{R}^+)$ and $x \in \mathbf{R}^+$, we have $\mathcal{M}'_\theta h(x) = \mathcal{M}_{T_{\theta,x}} h(x)$, where $T_{\theta,x} = \{x\} \cup U_{\theta,x}$, and $U_{\theta,x}$ is the union of all open intervals $A \subset \mathbf{R}^+$ such that (1) $x \in A$, and (2) θ is affine on A.

- The directional derivative $\mathcal{M}'_\theta$ is linear if and only if θ is strictly concave, in which case $\mathcal{M}'_\theta$ is the identity on $C_0(\mathbf{R}^+)$.
- The result above is from Beare and Fang (2017). Similar result appears in Beare and Moon (2015) but with θ a continuously differentiable concave CDF on $[0, 1]$.


1
θ
0
1


1
θ
0
1
h


1
θ
0
1
$\mathcal{M}'_R h$
h


1
θ
0
$\mathcal{M}'_R(-h)$
1
$-h$

Application #1: Testing for a monotone density ratio

- Let F and G be continuous CDFs with common support.
- Let $R = F \circ G^{-1}$, the associated ordinal dominance curve (ODC). Assume R is continuously differentiable on $[0, 1]$.
- We want to test the null hypothesis that R is concave.
- When F and G admit PDFs f and g, concavity of R is equivalent to nonincreasing-ness of the ratio f/g.
- We observe independent iid samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ drawn from F and G, and compute the empirical CDFs F_n and G_n.
- The empirical ODC is $R_n = F_n \circ G_n^{-1}$.
- Carolan and Tebbs (2005) suggest basing a test of the concavity of R on the test statistic
$$T_n = \sqrt{n}\|\mathcal{M}R_n - R_n\|_p,$$
with $p \in [1, \infty]$.
- We can use the delta method to study the behavior of their statistic.

Application #1: Testing for a monotone density ratio

- A standard application of the delta method reveals that

$$\sqrt{n}(R_n - R) \rightsquigarrow G_R := B_1 \circ R + R' \cdot B_2 \quad \text{in } \ell^\infty([0,1]),$$

  where B_1 and B_2 are independent standard Brownian bridges.
- Let $\mathcal{D} = \mathcal{M} - \mathcal{I}$, where $\mathcal{I}$ is the identity.
- When R is concave, another application of the delta method gives

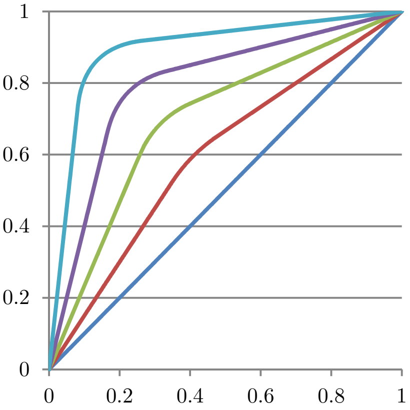
$$\sqrt{n}\mathcal{D}R_n = \sqrt{n}(\mathcal{D}R_n - \mathcal{D}R) \rightsquigarrow \mathcal{D}'_R G_R = \mathcal{M}'_R G_R - G_R \quad \text{in } \ell^\infty([0,1]).$$

- Thus from the continuous mapping theorem we have

$$T_n = \sqrt{n}\|\mathcal{D}R_n\|_p \rightsquigarrow \|\mathcal{D}'_R G_R\|_p \quad \text{in } \mathbf{R}.$$

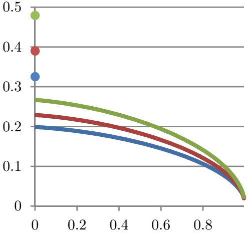
- What does $\|\mathcal{D}'_R G_R\|_p$ look like?

Application #1: Testing for a monotone density ratio

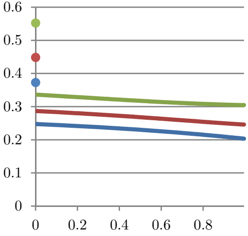


Ordinal dominance curves used to compute quartiles of $\|\mathcal{D}'_R G_R\|_p$. We plot the curves corresponding to $\delta = 0, 0.2, 0.4, 0.6, 0.8$. The curves shift upward as δ increases.

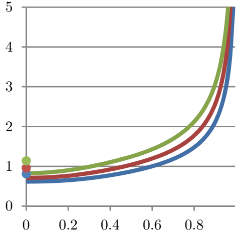
Application #1: Testing for a monotone density ratio



(a) $p = 1$



(b) $p = 2$



(c) $p = \infty$

Quartiles of $\|\mathcal{D}'_R G_R\|_p$, with $p = 1, 2, \infty$. The horizontal axes track the parameter δ indexing R.

Application #1: Testing for a monotone density ratio

$$\|\mathcal{D}_R' G_R\|_p = \left(\sum_{k \in K} \left(\lambda h_k d_k^{2/p} + (1-\lambda) h_k^2 d_k^{(2-p)/p} \right)^{p/2} \|\mathcal{D} B_k\|_p^p \right)^{1/p}$$

Application #1: Testing for a monotone density ratio

- We show that:
  - When $p \leq 2$, the distribution of $\|\mathcal{D}'_R G_R\|_p$ is maximal (in the sense of FOSD) when R is the 45° line. In this case it may be written as $\|\sqrt{2}\mathcal{D}B\|_p$.
  - When $p > 2$, the distribution of $\|\mathcal{D}'_R G_R\|_p$ diverges to infinity along a suitably chosen sequence of concave R's.
- Conclusion:
  - Don't use $p > 2$. Use $p \leq 2$.
  - Reject concavity if T_n exceeds the $(1 - \alpha)$-quantile of $\|\sqrt{2}\mathcal{D}B\|_p$.

Application #2: Bootstrapping the LCM of an empirical CDF

- Let $F : \mathbf{R}^+ \to \mathbf{R}$ be a concave CDF.
- Let $\mathbb{F}_n : \mathbf{R}^+ \to \mathbf{R}$ be the empirical CDF of n iid draws from F.
- Let $\mathbb{G} = B \circ F$, so that
$$\sqrt{n}(\mathbb{F}_n - F) \rightsquigarrow \mathbb{G}.$$
- We can obtain the weak limit of $\sqrt{n}(\mathcal{M}\mathbb{F}_n - F)$ by applying the delta method:
$$\sqrt{n}(\mathcal{M}\mathbb{F}_n - F) \rightsquigarrow \mathcal{M}'_F\mathbb{G}.$$
- This extends a similar result of Carolan (2002).

Application #2: Bootstrapping the LCM of an empirical CDF

- Now let $\mathbb{F}_n^*$ be a bootstrap version of $\mathbb{F}_n$, so that conditional on the data we have

$$\sqrt{n}(\mathbb{F}_n^* - \mathbb{F}_n) \rightsquigarrow \mathbb{G}.$$

- It would be nice if we could apply the delta method for the bootstrap to obtain

$$\sqrt{n}(\mathcal{M}\mathbb{F}_n^* - \mathcal{M}\mathbb{F}_n) \rightsquigarrow \mathcal{M}_F'\mathbb{G}$$

  conditional on the data. However, Dümbgen (1993) showed that the delta method cannot be applied in this way when we only have directional differentiability.

- Instead, we have (unconditionally)

$$\sqrt{n}(\mathcal{M}\mathbb{F}_n^* - \mathcal{M}\mathbb{F}_n) \rightsquigarrow \mathcal{M}_F'(\mathbb{G} + \mathbb{G}') - \mathcal{M}_F'(\mathbb{G}'),$$

  where $\mathbb{G}'$ is an independent copy of $\mathbb{G}$. Thus the bootstrap fails.

Bootstrapping the least concave majorant of a distribution function

- Consistent inference may be achieved by applying the rescaled bootstrap proposed (but not recommended) by Dümbgen (1993).
- Let $\hat{\mathcal{M}}'_n$ be given by

$$\hat{\mathcal{M}}'_n h = \frac{\mathcal{M}(\mathbb{F}_n + t_n h) - \mathcal{M}(\mathbb{F}_n)}{t_n}, \quad h \in \ell^\infty(\mathbf{R}^+),$$

  where $t_n \to 0$ and $\sqrt{n} t_n \to \infty$.
- Then $\hat{\mathcal{M}}'_n(\mathbb{F}^*_n - \mathbb{F}_n) \rightsquigarrow \mathcal{M}'_F \mathbb{G}$ conditional on the data.
- Finite sample performance seems to be quite poor.
- The rescaled bootstrap has been rediscovered in econometrics and is going by the name “numerical bootstrap”.

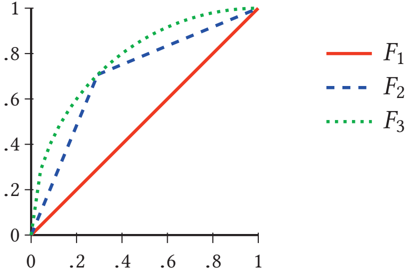
Application #2: Bootstrapping the LCM of an empirical CDF

- Suppose that instead we resample from the Grenander estimator. That is, let $\mathbb{F}_n^*$ be the empirical CDF of n draws from $\mathcal{M}\mathbb{F}_n$.
- Results of Sen, Banerjee and Woodroofe (2010) suggest that this is a bad idea.
- Indeed, we apply results of Kosorok (2008) to show that (unconditionally)

$$\sqrt{n}(\mathcal{M}\mathbb{F}_n^* - \mathcal{M}\mathbb{F}_n) \rightsquigarrow \mathcal{M}_F'(\mathbb{G} + \mathcal{M}_F'(\mathbb{G}')) - \mathcal{M}_F'(\mathbb{G}'),$$

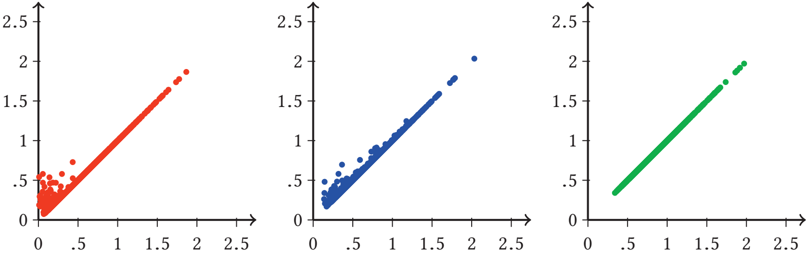
  where $\mathbb{G}'$ is an independent copy of $\mathbb{G}$. We do not achieve the desired limit $\mathcal{M}_F'\mathbb{G}$.
- **Exception**: Bootstrapping from the Grenander estimator can be used to approximate the upper quantiles of $\|\mathcal{M}_F'\mathbb{G}\|_\infty$. Numerically, the upper quantiles of $\|\mathcal{M}_F'\mathbb{G}\|_\infty$ and $\|\mathcal{M}_F'(\mathbb{G} + \mathcal{M}_F'(\mathbb{G}')) - \mathcal{M}_F'(\mathbb{G}')\|_\infty$ appear to be identical.

Application #2: Bootstrapping the LCM of an empirical CDF



Distribution functions used in numerical simulations.

Application #2: Bootstrapping the LCM of an empirical CDF



(a) F_1 (b) F_2 (c) F_3

Scatterplots of $\|\mathcal{M}'_F(\mathbb{G})\|_\infty$ versus $\|\mathcal{M}'_F(\mathbb{G} + \mathcal{M}'_F(\mathbb{G}')) - \mathcal{M}'_F(\mathbb{G}')\|_\infty$.

Application #3: Isotonic regression

- Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be n iid pairs of random variables satisfying

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, \ldots, n,$$

  where $m : \mathbf{R} \to \mathbf{R}$ is nondecreasing and $\varepsilon_1, \ldots, \varepsilon_n$ are iid centered random variables independent of $X_1, \ldots, X_n$.
- Let the X_i's arranged in ascending order be denoted by $X_{(1)}, \ldots, X_{(n)}$, and the corresponding Y_i's and ε_i's by $Y_{(1)}, \ldots, Y_{(n)}$ and $\varepsilon_{(1)}, \ldots, \varepsilon_{(n)}$.
- The isotonic regression estimator of m can be obtained from the left-derivative of the greatest convex minorant (GCM) of the cumulative sum diagram (CSD).
- Let's formulate the CSD as an element of $\ell^\infty([0,1])$:

$$\mathbb{S}_n(u) = \frac{1}{n} \sum_{i=0}^{[nu]} Y_{(i)} + \frac{nu - [nu]}{n} Y_{([nu]+1)}.$$

  Here, we set $Y_{(0)} = Y_{(n+1)} = 0$.
- Can we use the delta method to study the behavior of the GCM of $\mathbb{S}_n$?

Application #3: Isotonic regression

- Under some regularity conditions we can show that

$$\begin{aligned}\mathbb{S}_n(u) &= \frac{1}{n}\sum_{i=0}^{[nu]} m(X_{(i)}) + \frac{1}{n}\sum_{i=0}^{[nu]} \varepsilon_{(i)} + o_P(n^{-1/2}) \\ &= \int_0^u m(\mathbb{Q}_n(t))\mathrm{d}t + \frac{1}{n}\sum_{i=0}^{[nu]} \varepsilon_{(i)} + o_P(n^{-1/2}),\end{aligned}$$

  where $\mathbb{Q}_n$ is the empirical quantile function of the X_i's.
- This suggests setting

$$S(u) = \int_0^u m(Q(t))\mathrm{d}t.$$

- It is now straightforward to show using the delta method that

$$\sqrt{n}\,(\mathbb{S}_n(u) - S(u)) \rightsquigarrow -\int_0^u \frac{m'(Q(t))}{f(Q(t))}\mathbb{B}(t)\mathrm{d}t + \sigma\mathbb{W}(u),$$

  where $\mathbb{B}$ is a Brownian bridge and $\mathbb{W}$ an independent Brownian motion. Also, f is the PDF of the X_i's and σ^2 is the variance of the ε_i's.

Application #3: Isotonic regression

- Let $\hat{\mathbb{S}}_n = -\mathcal{M}(-\mathbb{S}_n)$, the GCM of $\mathbb{S}_n$.
- Another application of the delta method can be used to determine the weak limit of $\sqrt{n}\left(\hat{\mathbb{S}}_n - S\right)$ in terms of the directional derivative of $\mathcal{M}$.
- In particular, if m is flat then we have

$$\sqrt{n}\left(\hat{\mathbb{S}}_n - S\right) \rightsquigarrow -\mathcal{M}(-\sigma\mathbb{W}).$$

- Similarly, if m is flat then we can use the delta method to show that

$$\sqrt{n}\left(\hat{\mathbb{S}}_n - \mathbb{S}_n\right) \rightsquigarrow \sigma\mathcal{D}\mathbb{W},$$

  whereas if m is strictly increasing then we have

$$\sqrt{n}\left(\hat{\mathbb{S}}_n - \mathbb{S}_n\right) \rightsquigarrow 0.$$

- This suggests the possibility of testing the null hypothesis that m is flat by comparing a statistic $T_n = \sqrt{n}\|\hat{\mathbb{S}}_n - \mathbb{S}_n\|_p/\hat{\sigma}$ to the lower quantiles of $\|\mathcal{D}\mathbb{W}\|_p$.

Key references

[1] Beare, B. K. and Moon, J. -M. (2015). Nonparametric tests of density ratio ordering. *Econometric Theory*, 31(3):471-492.

[2] Beare, B. K. and Fang, Z. (2017). Weak convergence of the least concave majorant of estimators for a concave distribution function. *Electronic Journal of Statistics*, 11(2):3841-3870.

[3] Carolan, C. A. (2002). The least concave majorant of the empirical distribution function. *Canadian Journal of Statistics*, 30(2):317-328.

[4] Carolan, C. A. and Tebbs, J. M. (2005). Nonparametric tests for and against likelihood ratio ordering in the two sample problem. *Biometrika*, 92(1):159-171.

[5] Dümbgen, L. (1993). On nondifferentiable functions and the bootstrap. *Probability Theory and Related Fields*, 95(1):125-140.

[6] Fang, Z. and Santos, A. (2016). Inference on directionally differentiable functions. arXiv:1404.3763v2 [math.ST].

[7] Sen, B., Banerjee, M. and Woodroofe, M. (2010). Inconsistency of bootstrap: the Grenander estimator. *Annals of Statistics*, 38(4):1953-1977.

[8] Shapiro, A. (1990). On concepts of directional differentiability. *Journal of Optimization Theory and Applications*, 66(3):477-487.

[9] Shapiro, A. (1991). Asymptotic analysis of stochastic programs. *Annals of Operations Research*, 30(1):169-186.

[10] Tang, C. -F., Wang, D. and Tebbs, J. M. (2017). Nonparametric goodness-of-fit tests for uniform stochastic ordering. *Annals of Statistics*, 45(6):2565-2589.