Inference on distributions and quantiles using a finite-sample Dirichlet process

David M. Kaplan
University of Missouri

Matt Goldman
UC San Diego

Midwest Econometrics Group 2013
Outline

1. Dirichlet distribution and inference on distributions
2. Dirichlet process and inference on quantiles
3. Theoretical results and practical methods
4. Simulations
5. Conclusion
1 Dirichlet distribution and inference on distributions

2 Dirichlet process and inference on quantiles

3 Theoretical results and practical methods

4 Simulations

5 Conclusion
Wilks (1962)

- $Y_i \overset{iid}{\sim} F$, continuous $\implies F(Y_i) \overset{iid}{\sim} \text{Uniform}(0, 1)$
- $Y_{n:k}$: $k$th order statistic ($k$th-smallest value in sample); $F(Y_{n:k}) \overset{d}{=} U_{n:k}$
- $\{F(Y_{n:1}), \ldots, F(Y_{n:n})\} \sim \text{Dir}^*(1, \ldots, 1; 1)$
- $F(Y_{n:k}) \sim \beta(k, n + 1 - k)$
Exact inference on distributions

- Beta: $1 - \tilde{\alpha}$ CI for $F(\cdot)$ at each $Y_{n:k}$
- Dirichlet: which $\tilde{\alpha}$ makes $1 - \alpha$ uniform confidence band
- Hypothesis testing, incl. 2-sample/FOSD
- $(n$ tests controlling FWER$)$
- Advantages:
  - Exact, finite-sample coverage/type I error
  - While KS is insensitive to tails, and weighted KS overly sensitive to tails, Dirichlet allows exact relative pointwise type I error specification
  - Two-sample theory doesn’t require similar sample sizes
Pointwise type I error, n=20

- Dirichlet
- KS
- Weighted KS

Rejection probability

Order statistic

David M. Kaplan (U Missouri), Matt Goldman (UCSD)

Dirichlet-based distributional and quantile inference
Pointwise type I error, n=100

Rejection probability

Order statistic

Dirichlet  KS  weighted KS
Let \( n = 11 \), consider only \( k = 9 \)

- CI for CDF: take quantiles of \( F(Y_{11:9}) \sim \beta(9, 3) \)
- CI for quantile:

\[
P(F(Y_{11:9}) < 0.53) = 5\% = P(Y_{11:9} < F^{-1}(0.53)),
\]

so \( Y_{11:9} \) is endpoint for lower one-sided CI for 0.53-quantile

- Exact, finite-sample coverage
- But what if I care about the median instead?
1. Dirichlet distribution and inference on distributions

2. Dirichlet process and inference on quantiles

3. Theoretical results and practical methods

4. Simulations

5. Conclusion
Quantile inference via asymptotic normality

Ex: inference on median $Q_Y(0.5)$, using iid $\{Y_i\}_{i=1}^n$
Quantile inference via asymptotic normality

Ex: inference on median $Q_Y(0.5)$, using iid $\{Y_i\}_{i=1}^n$

$\sqrt{n}(\hat{Q} - Q_0) \overset{d}{\to} N(0, \sigma_Q^2)$
Dirichlet IDEAL Theory/methods Simulations

Normality Order statistics Dirichlet

Quantile inference via asymptotic normality

Ex: inference on median \( Q_Y(0.5) \), using iid \( \{Y_i\}_{i=1}^n \)

- \( \sqrt{n}(\hat{Q} - Q_0) \xrightarrow{d} N(0, \sigma_Q^2) \)
- 1-sided: \( (-\infty, \hat{Q} + 1.64\hat{\sigma}_Q/\sqrt{n}) \)
- Why do we do that?

\[
P(\hat{Q} + 1.64\sigma/\sqrt{n} < Q_0) = 0.05 = \alpha
\]

But: \( \sigma \) hard to estimate; asymptotic approx (worse in tails)
Normal Approximation for Uniform Sample 0.50-quantile, $n = 11$

- **Exact**
- **Normal approx**
Normal Approximation for Uniform Sample 0.58-quantile, $n = 11$

Density

Exact

Normal approx

$u$
Normal Approximation for Uniform Sample 0.67-quantile, $n = 11$

- **Exact**
- **Normal approx**

**Graph:**
- **Density**
- **u**

Data points:
- $0.0$ 1
- $0.2$ 2
- $0.4$ 3
- $0.6$ 4
- $0.8$ 5
- $1.0$ 6

Graph comparing exact and normal approximations for uniform sample quantiles with $n = 11$. The normal approximation closely follows the exact distribution.
Normal Approximation for Uniform Sample 0.75-quantile, $n = 11$

- **Exact**
- **Normal approx**

Graph showing the comparison between exact and normal approximations for the 0.75-quantile of a uniform sample with $n = 11$. The density of the data is plotted against $u$.
Normal Approximation for Uniform Sample 0.83-quantile, $n = 11$

- Exact
- Normal approx
Normal Approximation for Uniform Sample 0.92-quantile, $n = 11$

- Exact
- Normal approx
Quantile inference via order statistics

Ex: inference on median $Q_Y(0.5)$, using iid $\{Y_i\}^n_{i=1}$

- Order statistic: $Y_{n:k}$ is $k$th smallest out of $n$ in $\{Y_i\}^n_{i=1}$
- Approach: $(-\infty, Y_{n:k}]$ as lower one-sided CI; which $k$?
Quantile inference via order statistics

Ex: inference on median $Q_Y(0.5)$, using iid $\{Y_i\}_{i=1}^n$

- Order statistic: $Y_{n:k}$ is $k$th smallest out of $n$ in $\{Y_i\}_{i=1}^n$

- Approach: $(-\infty, Y_{n:k}]$ as lower one-sided CI; which $k$?

- Want: $\alpha = P(Y_{n:k} < F_Y^{-1}(0.5)) = P(U_{n:k} < 0.5)$

- Use: $U_i \sim \text{Unif}(0, 1)$
Endpoint selection using beta distribution

\[ \alpha = 0.05 = P(U_{n:k} < 0.5) \]

\[ U_{n:k} \sim \text{Beta}(k, n + 1 - k), \ k \in \{1, 2, \ldots, n\} \]
Endpoint selection using beta distribution

$$\alpha = 0.05 = P(U_{n:k} < 0.5)$$

$$U_{n:k} \sim \text{Beta}(k, n + 1 - k), \ k \in \{1, 2, \ldots, n\}$$

Determination of Upper Endpoint

n=11, p=0.5

$$P(U_{n:6} < 0.5) = 50\%$$

\[ k = 6 \]
Endpoint selection using beta distribution

\[ \alpha = 0.05 = P(U_{n:k} < 0.5) \]

\[ U_{n:k} \sim \text{Beta}(k, n + 1 - k), \quad k \in \{1, 2, \ldots, n\} \]

Determination of Upper Endpoint

\[ n=11, \quad p=0.5 \]

\[ P(U_{n:6} < 0.5) = 50\% \]

\[ k = 6 \quad (11.3\%) \]

\[ k = 8 \]
Endpoint selection using beta distribution

\[ \alpha = 0.05 = P(U_{n:k} < 0.5) \]

\[ U_{n:k} \sim \text{Beta}(k, n + 1 - k), \quad k \in \{1, 2, \ldots, n\} \]

Determination of Upper Endpoint

\[ n=11, \quad p=0.5 \]

\[ P(U_{n:6} < 0.5) = 50\% \]

\[ k = 6 \quad 11.3\% \]

\[ k = 8 \]

\[ k = 9 \quad 3.3\% \]
Endpoint selection using beta distribution

\[ \alpha = 0.05 = P(U_{n:k} < 0.5) \]

\[ U_{n:k} \sim \text{Beta}(k, n + 1 - k), \quad k \in \{1, 2, \ldots, n\} \quad \Rightarrow \quad k \in [1, n] \subset \mathbb{R} \]

Determination of Upper Endpoint

\[ n=11, \ p=0.5, \ \alpha=0.05 \]

\[ P(U_{n:6} < 0.5) = 50\% \]

\[ k = 6 \quad 11.3\% \]

\[ k = 8 \]

\[ k = 9 \quad 3.3\% \]
Endpoint selection using beta distribution

\[ \alpha = 0.05 = P(U_{n:k} < 0.5) \]

\[ U_{n:k} \sim \text{Beta}(k, n + 1 - k), \; k \in \{1, 2, \ldots, n\} \]

Determination of Upper Endpoint

\[ n=11, \; p=0.5 \]

\[ k=6 \]

\[ P(U_{n:6} < 0.5) = 50\% \]

\[ k=8 \]

\[ 11.3\% \]

\[ k=9 \]

\[ 3.3\% \]

Determination of Upper Endpoint

\[ n=11, \; p=0.5, \; \alpha=0.05 \]

\[ k=8.69 \]

\[ \alpha = 5\% \]
\[ \alpha = P(U_{n:k} < 0.5), \quad U_{n:k}^{I} \sim \text{Beta}(k, n + 1 - k) \]
Fractional order statistics

\[ \alpha = P(U_{n:k} < 0.5), \quad U_{n:k}^I \sim \text{Beta}(k, n + 1 - k) \]

Connecting back to reality, \( k \in [1, n] \subset \mathbb{R} \):

\[ (\text{observed}) \quad U_{n:k}^L = (1 - \epsilon)U_{n:[k]} + \epsilon U_{n:[k]+1}, \quad \epsilon \equiv k - \lfloor k \rfloor \]
Fractional order statistics

\[ \alpha = P(U_{n:k} < 0.5), \ U^I_{n:k} \sim \text{Beta}(k, n + 1 - k) \]

Connecting back to reality, \( k \in [1, n] \subset \mathbb{R} \):

(observed) \( U^L_{n:k} = (1 - \epsilon)U_{n:[k]} + \epsilon U_{n:[k]+1}, \quad \epsilon \equiv k - [k] \)

(Jones 2002) \( U^I_{n:k} \overset{d}{=} (1 - C\epsilon)U_{n:[k]} + C\epsilon U_{n:[k]+1}, \quad C\epsilon \sim \text{Beta}(\epsilon, 1 - \epsilon) \)
Recap:

- Endpoint is $Y_{n:k}$ (instead of $\hat{Q} + 1.64\hat{\sigma}/\sqrt{n}$)
- Determine exact, fractional $k$
- Approx $Y^I_{n:k}$ by $Y^L_{n:k}$, Interpolated Dual of Exact Analytic $L$-statistic (IDEAL): $O(n^{-1})$ coverage probability error
Sample Realizations from
`Ideal' Uniform Fractional Order Statistic Process

\[ \hat{Q}_1(u) \]

- \[ n=5 \]
- \[ n=100 \]
1. Dirichlet distribution and inference on distributions

2. Dirichlet process and inference on quantiles

3. Theoretical results and practical methods

4. Simulations

5. Conclusion
Assumptions

Assumption (PDF)

For each quantile $u_j$, the PDF $f(\cdot)$ satisfies (i) $f(F^{-1}(u_j)) > 0$; (ii) $f''(\cdot)$ is continuous in some neighborhood of $F^{-1}(u_j)$, i.e. $f \in C^2(U_\delta(F^{-1}(u_j)))$ with $U_\delta(x)$ denoting some $\delta$-neighborhood of point $x \in \mathbb{R}$.

Assumption (sampling)

Independent samples $\{X_i\}_{i=1}^{n_x}$ and $\{Y_i\}_{i=1}^{n_y}$ are drawn iid from respective CDFs $F_X$ and $F_Y$. Respective quantile functions are denoted $Q_X(\cdot) \equiv F_X^{-1}(\cdot)$ and $Q_Y(\cdot) \equiv F_Y^{-1}(\cdot)$.

Assumption (sample sizes)

Sample sizes grow as $\lim_{n_x \to \infty} \sqrt{\frac{n_y}{n_x}} \equiv \mu > 0$. 

Theorem: fractional order statistic processes

Wherever \( PDF \geq \delta \),

\[
\sup_{k \in (1, n)} |X_{n:k}^I - X_{n:k}^L| = O_p(n^{-1} \lfloor \log n \rfloor). 
\]
Lemma (Dirichlet PDF approximation)

Basically, precise approximation of Dirichlet PDF and derivative at values that for our purposes (i.e., for quantile inference) are drawn with probability quickly approaching one.
Theorem (CDF approximation)

Let $L^I = \text{lin. com. of idealized fractional order statistics}$,

$$L^L = \text{lin. com. of linearly interpolated fractional order statistics}.$$

(i) For samples $\{X_i\}_{i=1}^n$ with $X_i \overset{iid}{\sim} F$, and for a given constant $K$,

$$P \left( L^L < K \right) - P \left( L^I < K \right)$$

$$= C_1(K, n, \ldots)n^{-1} + O(n^{-3/2}\log(n)).$$

(ii) For samples $\{X_i\}_{i=1}^n$ with $X_i \overset{iid}{\sim} F$,

$$\sup_{K \in \mathbb{R}} \left[ P \left( L^L < K \right) - P \left( L^I < K \right) \right]$$

$$= C_2(n, \ldots)n^{-1} + O(n^{-3/2}\log(n))$$
Theorem (CDF approximation)

Let $L^I = \text{lin. com. of idealized fractional order statistics}$,

$$L^L = \text{lin. com. of linearly interpolated fractional order statistics.}$$

(iii) Given independent samples $\{X_i\}_{i=1}^{n_X}$ and $\{Y_i\}_{i=1}^{n_Y}$ with $X_i \overset{iid}{\sim} F_X$ and $Y_i \overset{iid}{\sim} F_Y$, with $n_x \propto n_y \propto n$,

$$\sup_{K \in \mathbb{R}} \left[ P(L^L_X + L^L_Y < K) - P(L^I_X + L^I_Y < K) \right] = O(n^{-1})$$
CPE: single quantile

Lower one-sided coverage probability is

\[ P\left( X_{n:k}^L(\alpha) > Q(u) \right) \]

Thm 3 \quad P\left( X_{n:k}^I(\alpha) > Q(u) \right) + \frac{\varepsilon_h (1 - \varepsilon_h) z_{1-\alpha} \exp\{-z_{1-\alpha}^2/2\}}{u^h(\alpha) (1 - u^h(\alpha)) \sqrt{2\pi}} n^{-1}

+ O(n^{-3/2} \log(n))

\[ P\left( X_{n:k}^I(\alpha) > Q(u) \right) = P\left( U_{n:k}^I(\alpha) > u \right) = 1 - \alpha. \]

\[ \Rightarrow \text{one- or two-sided CI: CPE is } O(n^{-1}) \]

\[ \Rightarrow \text{calibrated CI: CPE is } O(n^{-3/2} \log(n)) \]
CPE: quantile treatment effect

Coverage probability (lower one-sided CI)

\[
P\left\{ Y_{n:m(\alpha)}^L - X_{n:k(\alpha)}^L > Q_Y(p) - Q_X(p) \right\}
\]

\[
= P\left\{ Y_{n:m(\alpha)}^I - X_{n:k(\alpha)}^I > Q_Y(p) - Q_X(p) \right\} + O(n^{-1})
\]

\[
= P\left\{ F_Y^{-1}\left( U_{n:m(\alpha)}^{I,1} \right) - F_Y^{-1}(p) - \left[ F_X^{-1}\left( U_{n:k(\alpha)}^{I,2} \right) - F_X^{-1}(p) \right] > 0 \right\} + O(n^{-1})
\]

\[
= P\left\{ \frac{U_{n:m(\alpha)}^{I,1} - p}{f_Y(F_Y^{-1}(p))} - \frac{U_{n:k(\alpha)}^{I,2} - p}{f_X(F_X^{-1}(p))} > 0 \right\} + T + O(n^{-1})
\]

\[
= P\left\{ \gamma\left( U_{n:m(\alpha)}^{I,1} - p \right) > U_{n:k(\alpha)}^{I,2} - p \right\} + T + O(n^{-1})
\]

\[
= 1 - \alpha + C + T + O(n^{-1})
\]

C: conditioning on \( \hat{\gamma} \); T: Taylor remainder; \( O(n^{-1}) \): interpolation
CPE: joint, linear combinations

Linear combination: object of interest is $\sum_{j=1}^{J} \psi_j Q(u_j)$. CPE is same as for QTE: one-sided $O(n^{-1/2})$, two-sided $O(n^{-2/3})$

Joint: object of interest is $(Q(u_1), \ldots, Q(u_J))$. One- or two-sided CPE is $O(n^{-1})$

<table>
<thead>
<tr>
<th>Setting</th>
<th>Type</th>
<th>CPE (two-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-sample</td>
<td>single quantile</td>
<td>$O(n^{-3/2} \log(n))$</td>
</tr>
<tr>
<td>1-sample</td>
<td>joint (multiple quantiles)</td>
<td>$O(n^{-1})$</td>
</tr>
<tr>
<td>1-sample</td>
<td>linear combination</td>
<td>$O(n^{-2/3})$</td>
</tr>
<tr>
<td>2-sample</td>
<td>treatment effect</td>
<td>$O(n^{-2/3})$</td>
</tr>
</tbody>
</table>

\[
\text{CPE} \equiv P(\theta_0 \in \text{CI}) - (1 - \alpha)
\]
Nonparametric conditional model
Nonparametric conditional model
Setup

- Object of interest: $Q_{Y|X}(p; X = x_0)$
- Observables: $(X_i, Y_i) \in \mathbb{R}^d \times \mathbb{R}$
  - Discrete $X$ easily added: no bias

**Definition (effective sample)**

- Effective sample window: $C_h \equiv \{x : \|x - x_0\| \leq h\}$
- Effective sample: $\{Y_i : X_i \in C_h\}$
- Effective sample size: $N_n \equiv \#\{Y_i : X_i \in C_h\}$
Single quantile: pointwise, joint over multiple $x_0$
Joint over multiple quantiles

Pointwise (by x0) IDEAL 95% CIs for conditional joint upper and lower quartiles

Data

Upper quartile CI

Lower quartile CI
Linear combinations (interquartile range)

Pointwise IDEAL 95% CIs for conditional interquartile range (IQR)

- True IQR
- IDEAL CI
Conditional quantile treatment effects

IDEAL 95% CI (interpolated for visual ease) for conditional median treatment effect

- Data (treated)
- Data (control)
- Pointwise
- Joint
Inference on $Q_{Y|X}(p; x_0)$ using $Y_i$ in shrinking $C_h$

- Use 1-sample IDEAL method on effective sample; CPE in terms of $N_n$
- Derive nonparametric bias from using $X_i \neq x_0$
- Optimal bandwidth; overall CPE
Conditional: assumptions

(1) cts \((X_i, Y_i) \in \mathbb{R}^d \times \mathbb{R}, Y_i = Q_{Y|X}(p; X_i) + \epsilon_i, \epsilon_i \perp \epsilon_j | \{X_k\}_{k=1}^n,(\epsilon_i | X_i = x) \sim F_{\epsilon|X=x}

Assumption (bias)

(2) \(f_X(x_0) > 0; f_X(\cdot) \) has smoothness \(s_X = k_X + \gamma_X > 1 \) near \(x_0\)
(3) \(\forall u \) in neighborhood of \(p, Q_{Y|X}(u; \cdot) \) has smoothness \(s_Q = k_Q + \gamma_Q > 2\)
(4i) As \(n \to \infty, h \to 0\)

Assumption (IDEAL)

(4ii) As \(n \to \infty, nh^d/[\log(n)]^2 \to \infty (\implies N_n \overset{a.s.}{\to} \infty)\)
(5) \(f_{Y|X}(Q_{Y|X}(p; x_0); x_0) > 0 \) (unique cond. quantile)
(6) Smoothness: \(\forall y \) in neighborhood of \(Q_{Y|X}(u; x_0)\), \(\forall x \) in neighborhood of \(x_0\), \(f_{Y|X}(y; x) \in C^2 \) (wrt \(y\))
Bias: \( Q_{Y | C_h}(p) - Q_{Y | X}(p; x_0) \)

**Theorem (2)**

*Under Assumptions 2–4(i) and 6, \( \xi_p \equiv Q_{Y | X}(p; x_0) \):*

\[
\text{Bias} = -h^2 \frac{f_X(x_0) F^{(0,2)}_{Y | X}(\xi_p; x_0) + 2 f'_X(x_0) F^{(0,1)}_{Y | X}(\xi_p; x_0)}{6 f_X(x_0) f_{Y | X}(\xi_p; x_0)} + o(h^2)
\]
Optimal bandwidth and CPE rates (two-sided)

- CPE-optimal: \( h^* = \arg \min_h \text{CPE}(h) \)
- \( \text{CPE} = \text{CPE}_{\text{GK}} + \text{CPE}_{\text{Bias}} = O(N^{-1}) + O(N_n h^4) \)
- \( N_n \asymp n h^d \)
Optimal bandwidth and CPE rates (two-sided)

- CPE-optimal: $h^* = \arg\min_h \text{CPE}(h)$
- $\text{CPE} = \text{CPE}_{\text{GK}} + \text{CPE}_{\text{Bias}} = O(N_n^{-1}) + O(N_n h^4)$
- $N_n \asymp nh^d$

**Theorem (3)**

*Under Assumptions 1–6, two-sided CPE-optimal:*

$$h^* \asymp n^{-1/(2+d)} \quad \quad \text{CPE} = O(n^{-2/(2+d)})$$
Optimal bandwidth and CPE rates (two-sided)

- CPE-optimal: \( h^* = \arg \min_h \text{CPE}(h) \)
- \( \text{CPE} = \text{CPE}_{\text{GK}} + \text{CPE}_{\text{Bias}} = O(N_n^{-1}) + O(N_n h^4) \)
- \( N_n \asymp n h^d \)

**Theorem (3)**

*Under Assumptions 1–6, two-sided CPE-optimal:*

\[
    h^* \asymp n^{-1/(2+d)} \quad \text{CPE} = O(n^{-2/(2+d)})
\]

**Ex:** \( d = 1 \implies h^* \asymp n^{-1/3}, \text{CPE} = O(n^{-2/3}) \)
$O\left(n^{-2/3}\right)$
Optimal CPE comparison (two-sided)

- IDEAL: $\text{CPE} = O\left(n^{-2/(2+d)}\right)$
- Local polynomial/asy. normality: $\text{CPE} = O\left(n^{-\left[s_Q/(s_Q+d)\right]/2}\right)$
- $d = 1$ or $d = 2$: IDEAL better even if assume $s_Q = \infty$
- $d = 3$: IDEAL better unless assume $k_Q \geq 11$, 11th-degree local polynomial (364 terms)
- $d \to \infty$: IDEAL better unless $k_Q \geq 4$
**IDEAL CPE summary**

<table>
<thead>
<tr>
<th>Setting</th>
<th>Type</th>
<th>CPE (two-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-sample</td>
<td>single quantile</td>
<td>$O(n^{-3/2} \log(n))$</td>
</tr>
<tr>
<td>1-sample</td>
<td>joint (multiple quantiles)</td>
<td>$O(n^{-1})$</td>
</tr>
<tr>
<td>1-sample</td>
<td>linear combination</td>
<td>$O(n^{-2/3})$</td>
</tr>
<tr>
<td>2-sample</td>
<td>treatment effect</td>
<td>$O(n^{-2/3})$</td>
</tr>
<tr>
<td>conditional</td>
<td>single quantile</td>
<td>$O(n^{-2/(2+d)})$</td>
</tr>
<tr>
<td>conditional</td>
<td>joint</td>
<td>$O(n^{-2/(2+d)})$</td>
</tr>
<tr>
<td>conditional</td>
<td>linear combination</td>
<td>$O(n^{-8/(12+7d)})$</td>
</tr>
<tr>
<td>conditional</td>
<td>treatment effect</td>
<td>$O(n^{-8/(12+7d)})$</td>
</tr>
</tbody>
</table>

\[ \text{CPE} \equiv P(\theta_0 \in \text{CI}) - (1 - \alpha) \]
1. Dirichlet distribution and inference on distributions
2. Dirichlet process and inference on quantiles
3. Theoretical results and practical methods
4. Simulations
5. Conclusion
Unconditional QTE: power

Location difference of one unit between medians of $X$ and $Y$

$n_x = n_y = 25$, $F_X = F_Y$, $\alpha = 0.05$, $\rho = 0.5$

<table>
<thead>
<tr>
<th></th>
<th>N(0,1)</th>
<th>Logistic(0,1)</th>
<th>Exp(1)</th>
<th>LogN(0,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDEAL</td>
<td>0.79</td>
<td>0.39</td>
<td>0.91</td>
<td>0.75</td>
</tr>
<tr>
<td>K11</td>
<td>0.70</td>
<td>0.31</td>
<td>0.86</td>
<td>0.64</td>
</tr>
<tr>
<td>Hut07</td>
<td>0.75</td>
<td>0.38</td>
<td>0.87</td>
<td>0.57</td>
</tr>
<tr>
<td>BS-$t$ (99)</td>
<td>0.70</td>
<td>0.35</td>
<td>0.84</td>
<td>0.68</td>
</tr>
<tr>
<td>BS-$t$ (999)</td>
<td>0.71</td>
<td>0.36</td>
<td>0.86</td>
<td>0.68</td>
</tr>
<tr>
<td>Horowitz98</td>
<td>0.62</td>
<td>0.27</td>
<td>0.87</td>
<td>0.66</td>
</tr>
</tbody>
</table>
Conditional: other methods

- `quantreg::rqss` (nonparametric; est. asy. covariance matrix)
- `quantreg::rq` (parametric; bootstrap)
Koenker \texttt{rqss} simulations: iid errors (various), \( \sigma(X) = 0.2 \) or 0.2\((1 + X)\), median, \( n = 400 \), \( \alpha = 0.05 \)

\[
\sqrt{x(1-x)} \sin\left(2\pi\left(1 + 2^{-7/5}\right)/(x + 2^{-7/5})\right)
\]
Conditional: runtime

Runtimes (log-log)

Sample size

Runtime (minutes)

200 1,000 5,000 20,000 100,000

0.01 0.1 1 10 100

IDEAL

BS(99)
rqss
Pointwise CP: Gaussian, homoskedastic

Pointwise Coverage Probability

Nominal
IDEAL
rqss
rq(perc)
Pointwise CP: Cauchy, homoskedastic

Pointwise Coverage Probability

- Nominal
- IDEAL
- rqss
- rq(perc)
Pointwise CP: centered $\chi^2_3$, heteroskedastic

Pointwise Coverage Probability

- Nominal
- IDEAL
- rqss
- rq(perc)
Joint Hypotheses

0 no shift
+ shift = +0.1
Joint Hypotheses

0  no shift
+  shift = +0.1
-  shift = −0.1
Joint power curves: Gaussian, homoskedastic
Joint power curves: Cauchy, homoskedastic
Joint power curves: centered $\chi^2_3$, heteroskedastic

![Graph showing joint power curves with IDEAL, rqss, and rq(perc) methods.]
Pointwise power against $\pm 0.1$: Gaussian, homoskedastic

![Pointwise Power](image)

- **Power Rejection Probability (%)**
  - IDEAL
  - rqss

- **Deviation**: -0.1 or 0.1

- **Data Points**:
  - $0.2, 0.4, 0.6, 0.8$
  - $0, 20, 40, 60, 80, 100$

---

David M. Kaplan (U Missouri), Matt Goldman (UCSD)
Pointwise power against ±0.1: Cauchy, homoskedastic
Pointwise power against $\pm 0.1$: centered $\chi^2_3$, heteroskedastic
Conditional QTE: median

```
<table>
<thead>
<tr>
<th>CP</th>
<th>0.925</th>
<th>0.965</th>
<th>0.945</th>
<th>0.920</th>
<th>0.955</th>
<th>0.950</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>0.367</td>
<td>0.354</td>
<td>0.324</td>
<td>0.283</td>
<td>0.271</td>
<td>0.323</td>
</tr>
</tbody>
</table>
```
### Conditional quantile treatment effects

<table>
<thead>
<tr>
<th>$p$</th>
<th>0.04</th>
<th>0.224</th>
<th>0.408</th>
<th>0.592</th>
<th>0.776</th>
<th>0.96</th>
<th>Joint</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_0$ value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coverage Probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>0.938</td>
<td>0.924</td>
<td>0.938</td>
<td>0.940</td>
<td>0.948</td>
<td>0.960</td>
<td>0.952</td>
</tr>
<tr>
<td>0.50</td>
<td>0.926</td>
<td>0.962</td>
<td>0.964</td>
<td>0.952</td>
<td>0.944</td>
<td>0.952</td>
<td>0.946</td>
</tr>
<tr>
<td>0.25</td>
<td>0.948</td>
<td>0.960</td>
<td>0.948</td>
<td>0.930</td>
<td>0.922</td>
<td>0.924</td>
<td>0.932</td>
</tr>
<tr>
<td>Median Interval Length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>0.366</td>
<td>0.364</td>
<td>0.325</td>
<td>0.281</td>
<td>0.268</td>
<td>0.322</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>0.348</td>
<td>0.340</td>
<td>0.323</td>
<td>0.282</td>
<td>0.250</td>
<td>0.297</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.373</td>
<td>0.373</td>
<td>0.337</td>
<td>0.316</td>
<td>0.258</td>
<td>0.341</td>
<td></td>
</tr>
</tbody>
</table>

**Table:** CP and median interval length for IDEAL CIs for conditional QTEs; $1 - \alpha = 0.95$, $n = 400$ for both treatment and control samples, 500 replications.
1. Dirichlet distribution and inference on distributions

2. Dirichlet process and inference on quantiles

3. Theoretical results and practical methods

4. Simulations

5. Conclusion
Recap

- Wilks (1962) Dirichlet: exact inference on distributions
- Fractional order statistic theory: IDEAL inference on quantiles
- Methods: single quantile, joint, linear combinations, quantile treatment effects; unconditional, conditional
- Results: improved CPE in most common cases, robust
- R code, examples, papers on my website
Inference on distributions and quantiles using a finite-sample Dirichlet process

David M. Kaplan
University of Missouri

Matt Goldman
UC San Diego

Midwest Econometrics Group 2013