Seminar at Institute of Statistics
National University of Kaohsiung

2012, March 28

A goodness-of-fit test based on the multiplier Bootstrap with application to left-truncated data

by Takeshi Emura

Joint work with Dr. Yoshihiko Konno,
Japan Women's University

Outlines

- Goodness-of-fit test: Review
- Multiplier Bootstrap: Review

- Goodness-of-fit with application to left-truncated data
  1. Proposed method
  2. Simulation
  3. Data analysis

Goodness-of-fit: Review

- Parametric family:

$$\mathfrak{I} = \{ F_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \Theta \}, \qquad \Theta \subset \mathbf{R}^p$$

- Underlying CDF = F

- Testing

$$H_0 : F \in \mathfrak{I} \quad vs. \quad H_1 : F \notin \mathfrak{I}$$

based on data $X_1, \ldots, X_n \sim F$

•Under $H_0 : F \in \Im$,

there exists $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ such that $F = F_{\boldsymbol{\theta}}$.

$\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(X_1, \ldots, X_n)$ is consistent

$$\Rightarrow \begin{cases} F_{\hat{\boldsymbol{\theta}}}(x) \to F_{\boldsymbol{\theta}}(x) = F(x), \\ \hat{F}(x) \equiv \frac{1}{n} \sum_{i=1}^{n} I(X_i \le x) \to F(x) \end{cases}$$

Test Statistics

1) Kolmogorov-Smirnov statistic

$$K = \sup_{x \in R^k} | \hat{F}(x) - F_{\hat{\boldsymbol{\theta}}}(x) |,$$

2) Cramér-von Mises statistic

$$C = \int \{ \hat{F}(x) - F_{\hat{\boldsymbol{\theta}}}(x) \}^2 dF_{\hat{\boldsymbol{\theta}}}(x)$$

•The null distribution of :

$$K = \sup_{x \in R^k} | \hat{F}(x) - F_{\hat{\boldsymbol{\theta}}}(x) |$$

$$C = \int \{ \hat{F}(x) - F_{\hat{\boldsymbol{\theta}}}(x) \}^2 \, dF_{\hat{\boldsymbol{\theta}}}(x)$$

depend on both $\boldsymbol{\theta}$ and $\mathfrak{I} = \{ F_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \Theta \}$

1) $\mathfrak{I} = \{ N(\mu, \sigma^2) \mid \mu \in \mathbf{R}, \sigma^2 > 0 \}$

Table available in Lilliefors (1967 *JASA*)

2) $\mathfrak{I} = \{ Exp(\theta) \mid \theta > 0 \}$

Table available in Lilliefores (1969 *JASA*)

•*Parametric Bootstrap* has been suggested to approximate the null distribution of

$$K = \sup_{x \in R^k} | \hat{F}(x) - F_{\hat{\theta}}(x) |$$

$$C = \int \{ \hat{F}(x) - F_{\hat{\theta}}(x) \}^2 dF_{\hat{\theta}}(x)$$

Under cerntain regularity conditions, the parametric bootstrap is theoretically and numerically justified:

- Stute et al. (1993 *Metrika*): Continuous case
- Henze (1996 *Can. J. Stat*): Discrete case
- Genest & Remillard (2008 *Annales de Inst. Poincare*)
  : Easy-to-check regularity conditions
- p.279 of *Asymptotic Statistics* by van der Vaart (1998)
  : Empirical process approach

Algorithm:

Step 1: Generate $(X_1^{*(b)}, \ldots, X_n^{*(b)}) \sim F_{\hat{\boldsymbol{\theta}}}, \quad b = 1, \ldots, B$

Step 2: Calculate $\hat{\boldsymbol{\theta}}^{*(b)}$ and $\hat{F}^{*(b)}(x) = \frac{1}{n}\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{I}(X_1^{*(b)} \le x)$

from $(X_1^{*(b)}, \ldots, X_n^{*(b)})$

Step 3: Compute

$$C^{*(b)} = \int \{ \hat{F}^{*(b)}(x) - F_{\hat{\boldsymbol{\theta}}^{*(b)}}(x) \}^2 \, dF_{\hat{\boldsymbol{\theta}}^{*(b)}}(x)$$

Step 4: Reject $\mathrm{H}_0 : \mathfrak{I} = \{ F_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \Theta \}$ if

$$P\text{-value} = \frac{1}{\mathrm{B}} \sum_{\mathrm{b=1}}^{\mathrm{B}} \mathrm{I}(C^{*(b)} > C) < 0.05$$

- The parametric Bootstrap is very time consuming, especially when F is multivariate

- A tool to reduce the computational cost "Multiplier Bootstrap" or "weighted Bootstrap" become very popular in the recent literature
(Remillard & Scaillet 2009 JMVA; Kojadinovic & Yan 2011 Stat.& Computing ;Bucher & Dette 2010 Stat&Prob letters; Kojadinovic & Yan, 2012 manusc.)

- *I use 4 slide to explain what is "multiplier Bootstrap"*

What is multiplier Bootstrap?

- IID: $X_1, \ldots, X_n \sim F$
- Empirical Process:

$$\mathbf{G}_n(t) = \frac{1}{\sqrt{n}} \sum_i \{I(X_i \le t) - F(t)\}$$

How to approximate the distribution ?

- Re-sampling: $(X_1, \ldots, X_n) \to (X_1^*, \ldots, X_n^*)$

Bootstrap version:

$$\mathbf{G}_n^*(t) = \frac{1}{\sqrt{n}} \sum_i \{I(X_i^* \le t) - \hat{F}(t)\} \qquad \hat{F}(t) = \frac{1}{n} \sum_i I(X_i \le t)$$

What is multiplier Bootstrap?

- Multiplier representation

$$\mathbf{G}_n^*(t) = \frac{1}{\sqrt{n}} \sum_i \{I(X_i^* \le t) - \hat{F}(t)\}$$

$$= \frac{1}{\sqrt{n}} \sum_i Z_i \{I(X_i \le t) - \hat{F}(t)\}$$

$$\text{where} \quad (Z_1, \ldots, Z_n) \sim \text{Multinomial } \{n, (1/n, \ldots, 1/n)\}$$

Ref: Sec 3.6 of van der Vaart & Wellner (1996) *Weak convergence & empirical processes*

- Multiplier Bootstrap is a modification :

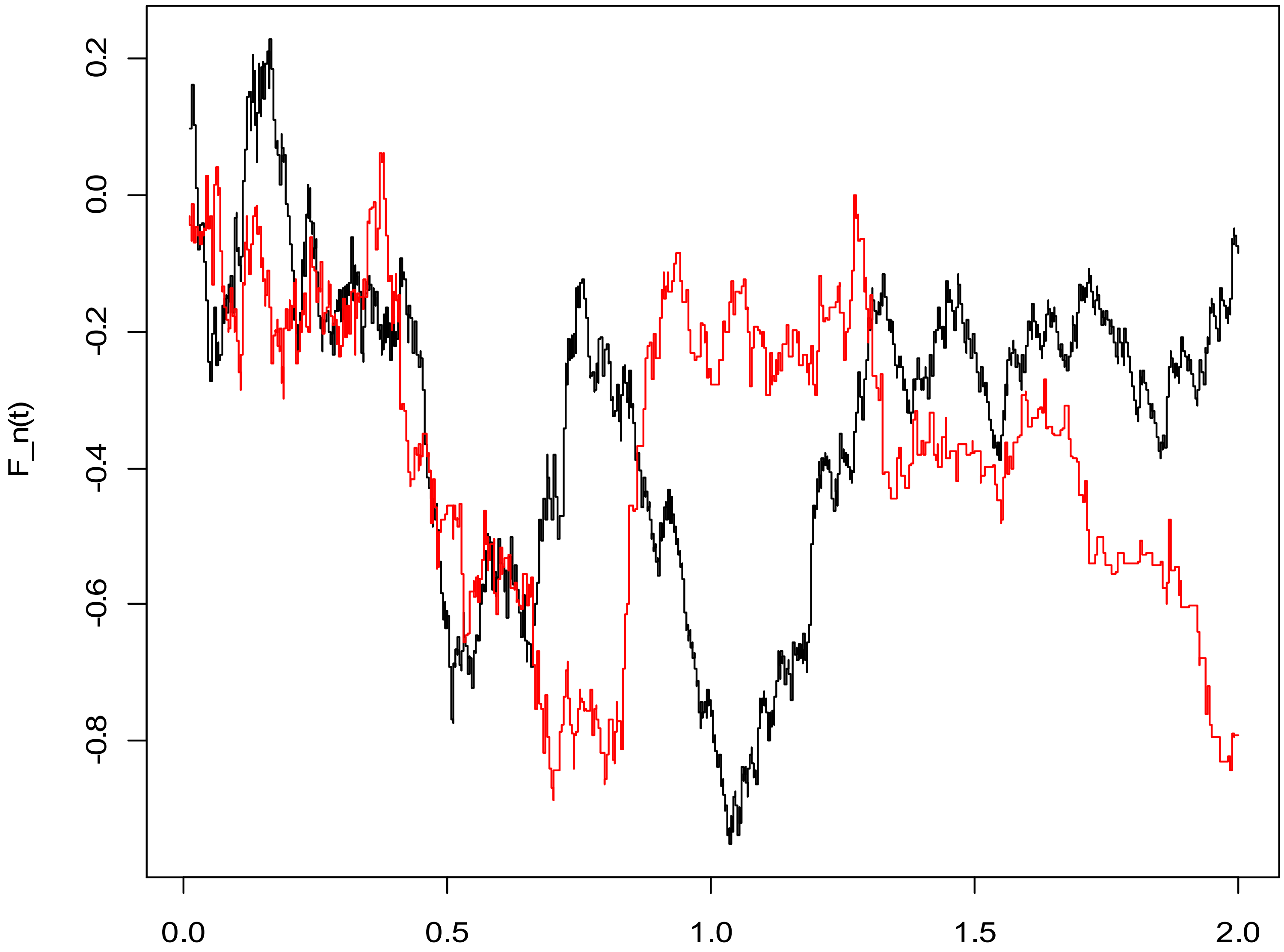
$$\mathbf{G}_n^{\mathrm{MP}}(t) = \frac{1}{\sqrt{n}} \sum_i Z_i \{I(X_i \le t) - \hat{F}(t)\}$$

$$(Z_1, \ldots, Z_n) \overset{iid}{\sim} \mathrm{Z}, \quad \mathrm{E}[\mathrm{Z}] = 0, \quad \mathrm{Var}(\mathrm{Z}) = 1$$

----- Original $\mathbf{G}_n(t)$

----- Multiplier Boot $\mathbf{G}_n^{\mathrm{MP}}(t)$

n=1000, X~Exp(1), Z~N(0,1)



$$\text{Multiplier central limit theorem :}$$

$$\mathbf{F}_n(t) \text{ and } \mathbf{F}_n^{\text{MP}}(t) \text{ has the same asymptotic distribution}$$

(Ref: Sec 2.9: Multiplier Central Limit Theorem in the book:
van der Vaart & Wellner (1996) *Weak convergence & empirical processes*)

- In principle, Multiplier Bootstrap is applicable to any statistics of the sum of iid terms:

$$\text{Statistic : } T = \frac{1}{\sqrt{n}} \sum_i U_i$$

$$\text{MP Bootstrap version : } T^* = \frac{1}{\sqrt{n}} \sum_i Z_i U_i$$

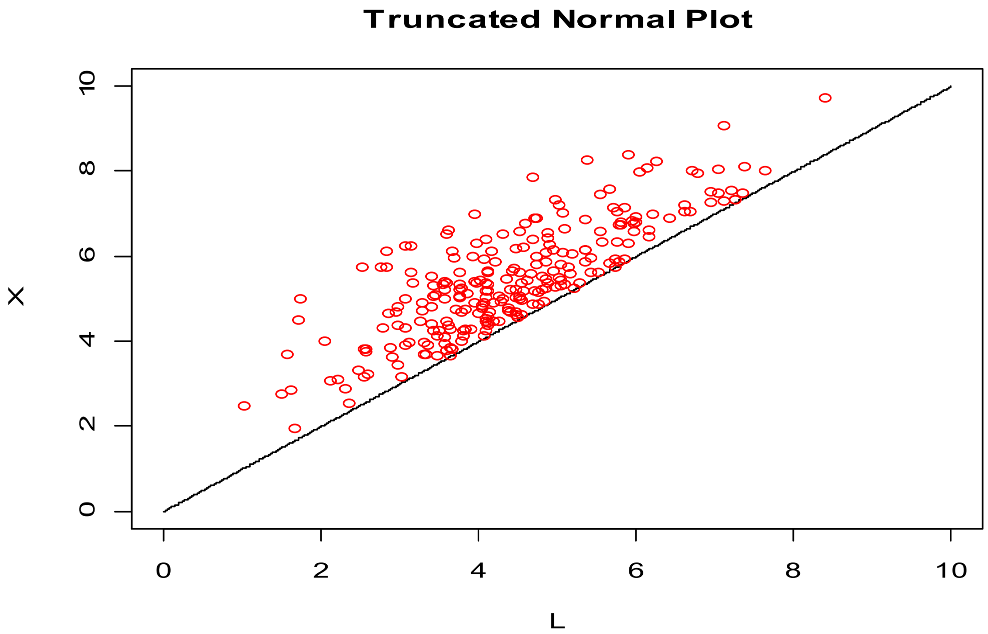
From now on, I focus on applications to left-truncated data

Left-Truncated data

- Vast literature: see

  Book of Klein & Moeschberger 2nd ed. 2003

$$\{(L_j, X_j)\,(j=1,\ldots,n)\} \quad \text{subject to } L_j \le X_j$$



L~N(5,2),
X~N(5,2),
Cov(L,X)=0.5

- Mathematical Set up

Population random variable; (L^O, X^O)

If $L^O > X^O$, nothing observed!

If $L^O \le X^O$, a pair $(L^O, X^O) = (L_j, X_j)$ is observed

Post-truncated data;

$$\{(L_j, X_j)\,(j = 1, \ldots, n)\} \quad \text{subject to } L_j \le X_j$$

$i.i.d.$ from

$$\Pr(L^O \in dl, X^O \in dx \mid L^O \le X^O) = \frac{f_{L^O X^O}(l, x)\mathbf{I}(l \le x)}{\Pr(L^O \le X^O)}$$

- Likelihood

$$L(\boldsymbol{\theta}) = c(\boldsymbol{\theta})^{-n} \prod_j f_{\boldsymbol{\theta}}(L_j, X_j),$$

$$\text{where} \quad \begin{cases} f_{\boldsymbol{\theta}}: \ \text{density for } (L^O, X^O) \\ c(\boldsymbol{\theta}) = \Pr(L^O \le X^O) = \iint_{l \le x} f_{\boldsymbol{\theta}}(l, x) dl dx. \end{cases}$$

- Maximum likelihood estimator

$$\hat{\boldsymbol{\theta}}: \ \frac{\partial}{\partial \boldsymbol{\theta}} \log L(\boldsymbol{\theta}) = \mathbf{0}$$

Consistency & Asymptotic normality follows when $c(\boldsymbol{\theta})$ and $f_{\boldsymbol{\theta}}$ are 3 times continuously differentiable (Emura & Konno 2012 CSDA)

Some useful bivariate models

Example 1: Bivariate *t* or bivariate normal

$$\boldsymbol{\theta}' = (\boldsymbol{\mu}' = (\mu_L, \mu_X), \sigma_L^2, \sigma_X^2, \sigma_{LX}, v)$$

$$f_{\boldsymbol{\theta}}(l, x) = \frac{(\sigma_L^2\sigma_X^2 - \sigma_{LX}^2)^{-1/2}\Gamma\{(v+2)/2\}}{\pi\,\Gamma(v/2)v}\left\{1+\frac{Q^2(l,x\,|\,\boldsymbol{\mu},\boldsymbol{\Sigma})}{v}\right\}^{-(v+2)/2}$$

$$c(\boldsymbol{\theta}) = \Pr(L^O \le X^O) = \Psi\left(\frac{\mu_X - \mu_L}{\sqrt{\sigma_X^2 + \sigma_L^2 - 2\sigma_{LX}}} : v\right), \quad \Psi(:v):\ \text{cdf of student-}t$$

Example 2: Bivariate Poisson (Holgate, 1964 *Biometrika*) :

$$f_{\boldsymbol{\theta}}(l,x) = e^{-\lambda_L-\lambda_X-\alpha}\sum_{w=0}^{l}\frac{\lambda_L^{l-w}\lambda_X^{x-w}\alpha^w}{(l-w)!(x-w)!w!}$$

$$c(\boldsymbol{\theta}) = \Pr(L^O \le X^O) = e^{-\lambda_L-\lambda_X}\sum_{l=0}^{\infty}\frac{\lambda_L^l}{l!}S_V(l) \qquad S_V(l) = \sum_{v=l}^{\infty}\frac{\lambda_X^v e^{-\lambda_X}}{v!}$$

Example 3: Zero-modified Poisson (Dietz and Böhning, 2000 *CSDA*)

$$f_{\boldsymbol{\theta}}(l,x) = p^l(1-p)^{1-l}\frac{(\lambda_X)^x e^{-\lambda_X}}{x!}, \qquad l = 0,1; \quad x = 0,1,2,\ldots$$

$$c(\boldsymbol{\theta}) = \Pr(L^O \le X^O) = 1 - pe^{-\lambda_X}$$

Let $\Im = \{ f_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \Theta \}$ be a given parametric family.

$$H_0 : f \in \Im \qquad \text{against} \qquad H_1 : f \notin \Im .$$

Goodness-of-fit

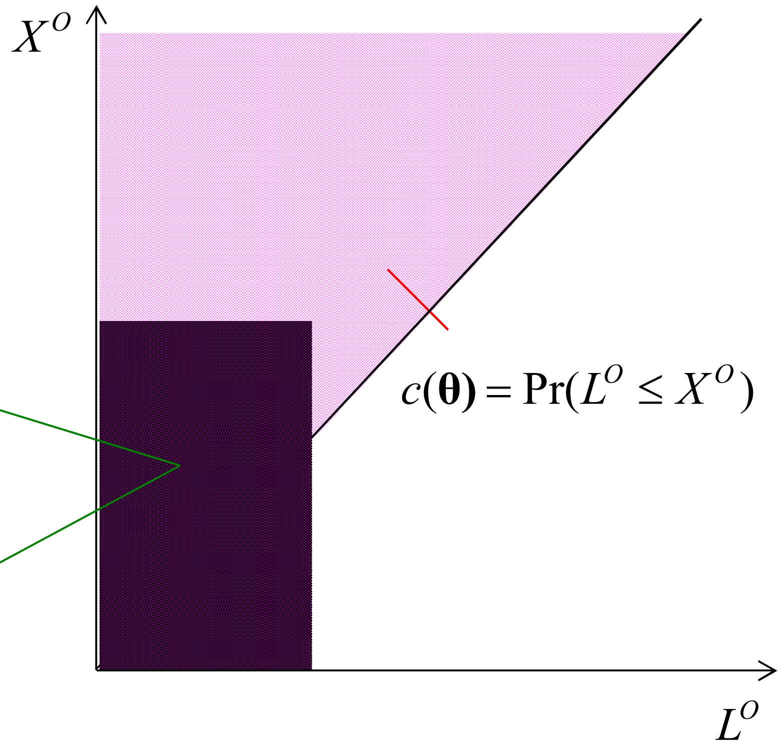
- <u>Cramér-von Mises statistic</u>

$$C = \iint_{l \le x} \{ \hat{F}(l,x) - F_{\hat{\boldsymbol{\theta}}}(l,x) / c(\hat{\boldsymbol{\theta}}) \}^2 d\hat{F}(l,x)$$

$$= \sum_j \{ \hat{F}(L_j, X_j) - F_{\hat{\boldsymbol{\theta}}}(L_j, X_j) / c(\hat{\boldsymbol{\theta}}) \}^2,$$

where $\hat{F}(l,x) = \sum_j \mathbf{I}(L_j \le l, X_j \le x) / n$

$$F_{\boldsymbol{\theta}}(l,x) = \begin{cases} \iint_{u \le l, u \le v \le x} f_{\boldsymbol{\theta}}(u,v)\,du\,dv & \text{continuous;} \\ \sum_{u \le l} \sum_{u \le v \le x} f_{\boldsymbol{\theta}}(u,v) & \text{discrete.} \end{cases}$$



* I will not use Kolmogorov-Smirnov statistic since it is computationally more demanding.

Some useful bivariate models

Example 1: Bivariate normal

$$f_{\boldsymbol{\theta}}(l,x) = \text{Bivariate normal density}$$

$$F_{\boldsymbol{\theta}}(l,x) = \int_{-\infty}^{(l-\mu_L)/\sigma_L} \left[\Phi\left\{ \frac{x-\mu_X-\sigma_{LX}s/\sigma_L}{\sqrt{\sigma_X^2-\sigma_{LX}^2/\sigma_L^2}} \right\} - \Phi\left\{ \frac{\mu_L+\sigma_L s-\mu_X-\sigma_{LX}s/\sigma_L}{\sqrt{\sigma_X^2-\sigma_{LX}^2/\sigma_L^2}} \right\} \right] \phi(s)ds$$

Example 2: Bivariate Poisson (Holgate, 1964) :

$$f_{\boldsymbol{\theta}}(l,x) = e^{-\lambda_L-\lambda_X-\alpha} \sum_{w=0}^{l} \frac{\lambda_L^{l-w}\lambda_X^{x-w}\alpha^w}{(l-w)!(x-w)!w!} \qquad F_{\boldsymbol{\theta}}(l,x) = e^{-\lambda_L-\lambda_X-\alpha} \sum_{u=0}^{l}\sum_{v=u}^{x}\sum_{w=0}^{u} \frac{\lambda_L^{u-w}\lambda_X^{v-w}\alpha^w}{(u-w)!(v-w)!w!}$$

Example 3: Zero-modified Poisson (Dietz and Böhning, 2000)

$$f_{\boldsymbol{\theta}}(l,x) = p^l(1-p)^{1-l}\frac{(\lambda_X)^x e^{-\lambda_X}}{x!}, \qquad l=0,1; \quad x=0,1,2,\ldots$$

$$F_{\boldsymbol{\theta}}(l,x) = \begin{cases} (1-p)e^{-\lambda_X}\sum_{u=0}^{x}(\lambda_X^x/u!) & \text{if } l=0 \;; \\ e^{-\lambda_X}\left(\sum_{u=0}^{x}\lambda_X^x/u! - p\right) & \text{if } l=1, \end{cases}$$

Parametric Bootstrap

Step 0: Calculate the statistic $C = \sum_i \{ \hat{F}(L_i, X_i) - F_{\hat{\boldsymbol{\theta}}}(L_i, X_i) / c(\hat{\boldsymbol{\theta}}) \}^2$.

Step 1: Generate $(L_j^{(b)}, X_j^{(b)})$ which follows the truncated distribution of $F_{\hat{\boldsymbol{\theta}}}(l, x)$,

subject to $L_j^{(b)} \le X_j^{(b)}$, for $b = 1, 2, \ldots, B,\ j = 1, 2, \ldots, n$.

Step 2: Calculate $\{ C^{*(b)}; b = 1, 2, \ldots, B \}$,

where $C^{*(b)} = \sum_i \{ \hat{F}^{(b)}(L_i^{(b)}, X_i^{(b)}) - F_{\hat{\boldsymbol{\theta}}^{(b)}}(L_i^{(b)}, X_i^{(b)}) / c(\hat{\boldsymbol{\theta}}^{(b)}) \}^2$ and where $\hat{F}^{(b)}(l, x)$ and $\hat{\boldsymbol{\theta}}^{(b)}$ are the empirical CDF and MLE based on $\{ (L_j^{(b)}, X_j^{(b)}); j = 1, 2, \ldots, n \}$.

Step 3: Reject H_0 at the 100α % significance level if $\sum_{b=1}^{B} \mathbf{I}(C^{*(b)} \ge C) / B < \alpha$.

Parametric Bootstrap

- Some simulation results under bivariate normal model (Emura & Konno, 2012a *Stat Papers*)

$$\begin{bmatrix} L^O \\ X^O \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_L \\ \mu_X \end{bmatrix} = \begin{bmatrix} 120 - 62.63 \\ 60.82 \end{bmatrix} \; \begin{bmatrix} \sigma_L^2 & \sigma_{LX} \\ \sigma_{LX} & \sigma_X^2 \end{bmatrix} = \begin{bmatrix} 19.64^2 & 19.64 \times 16.81 \rho_{LX} \\ 19.64 \times 16.81 \rho_{LX} & 16.81^2 \end{bmatrix} \right)$$

Table 1: Type I error rates of the the Cramér-von-Mises type goodness-of-fit test at level α based on 300 replications. The cut-off value is obtained by the parametric bootstrap-based procedure based on 1,000 resamplings.

	$\rho_{X^O Y^O}$				
	-0.70	-0.35	0.00	0.35	0.70
$\alpha = 0.10$	0.113	0.090	0.107	0.123	0.083
$\alpha = 0.05$	0.056	0.040	0.047	0.060	0.050
$\alpha = 0.01$	0.013	0.007	0.007	0.023	0.013

•Parametric Bootstrap is computationally very intensive, especially in **Step 2**:

Step 2: Calculate $\{ C^{*(b)}; b = 1, 2, \ldots, B \}$,

where $C^{*(b)} = \sum_i \{ \hat{F}^{(b)}(L_i^{(b)}, X_i^{(b)}) - F_{\hat{\boldsymbol{\theta}}^{(b)}}(L_i^{(b)}, X_i^{(b)}) / c(\hat{\boldsymbol{\theta}}^{(b)}) \}^2$ and where $\hat{F}^{(b)}(l, x)$ and $\hat{\boldsymbol{\theta}}^{(b)}$ are the empirical CDF and MLE based on $\{ (L_j^{(b)}, X_j^{(b)}); j = 1, 2, \ldots, n \}$.

• Maximizing likelihood functions *B* times often encounter the problem of local maxima

(especially serious when data are not from the null)

Alternatively, we apply the "Multiplier Bootstrap" to approximate the distribution of

$$C = \iint_{l \le x} \{ \hat{F}(l, x) - F_{\hat{\boldsymbol{\theta}}}(l, x) / c(\hat{\boldsymbol{\theta}}) \}^2 \, d\hat{F}(l, x)$$
$$= \sum_j \{ \hat{F}(L_j, X_j) - F_{\hat{\boldsymbol{\theta}}}(L_j, X_j) / c(\hat{\boldsymbol{\theta}}) \}^2$$

Motivation: Reduce the computational cost

Theorem 1 (Emura & Konno 2012 *CSDA*):

$$\sqrt{n}\{\hat{F}(l,x)-F_{\hat{\boldsymbol{\theta}}}(l,x)/c(\hat{\boldsymbol{\theta}})\}=\frac{1}{\sqrt{n}}\sum_j \hat{V}_j(l,x;\hat{\boldsymbol{\theta}})+o_P(1),$$

where $\hat{V}_j(l,x;\boldsymbol{\theta})=\mathbf{I}(L_j\le l, X_j\le x)-F_{\boldsymbol{\theta}}(l,x)/c(\boldsymbol{\theta})-\mathbf{g}'_{\boldsymbol{\theta}}(l,x)\{i_n(\boldsymbol{\theta})/n\}^{-1}\dot{l}_j(\boldsymbol{\theta})$,

$\mathbf{g}_{\boldsymbol{\theta}}(l,x)=c(\boldsymbol{\theta})^{-2}\{\dot{F}_{\boldsymbol{\theta}}(l,x)c(\boldsymbol{\theta})-F_{\boldsymbol{\theta}}(l,x)\dot{c}(\boldsymbol{\theta})\}$,

$\dot{F}_{\boldsymbol{\theta}}(l,x)=\partial F_{\boldsymbol{\theta}}(l,x)/\partial\boldsymbol{\theta}, \quad \dot{c}(\boldsymbol{\theta})=\partial c(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$

- Let Z_j be random variable with $\mathrm{E}(Z_j)=0, \; Var(Z_j)=1$

$$\frac{1}{\sqrt{n}}\sum_j Z_j\hat{V}_j(l,x;\hat{\boldsymbol{\theta}})$$

Conditional on data, only random quantities are

$$\{Z_j;\ j=1,2,\ldots,n\}$$

By the multiplier central limit theorem, the dsitribution of

$C = \sum_j \{ \hat{F}(L_j, X_j) - F_{\hat{\boldsymbol{\theta}}}(L_j, X_j) / c(\hat{\boldsymbol{\theta}}) \}^2$ is approximated by

$$\frac{1}{n^2} \sum_i \left[\sum_j Z_j \hat{V}_j (L_i, X_i; \hat{\boldsymbol{\theta}}) \right]^2 = \left\| \frac{\mathbf{Z}' \hat{\mathbf{V}}(\hat{\boldsymbol{\theta}})}{n} \right\|^2$$

Step 0: Calculate the statistic $C = \sum_i \{ \hat{F}(L_i, X_i) - F_{\hat{\boldsymbol{\theta}}}(L_i, X_i) / c(\hat{\boldsymbol{\theta}}) \}^2$ and matrix $\hat{\mathbf{V}}(\hat{\boldsymbol{\theta}})$.

Step 1: Generate $Z_j^{(b)} \sim N(0, 1)$; $b = 1, 2, \ldots, B$, $j = 1, 2, \ldots, n$.

Step 2: Calculate $\{ C^{(b)}; b = 1, 2, \ldots, B \}$,

where $C^{(b)} = \| (\mathbf{Z}^{(b)})' \hat{\mathbf{V}}(\hat{\boldsymbol{\theta}}) / n \|$ and $\mathbf{Z}^{(b)} = (Z_1^{(b)}, \ldots, Z_n^{(b)})'$.

Step 3: Reject H_0 at the 100α % significance level if $\sum_{b=1}^{B} \mathbf{I}(C^{(b)} \geq C) / B < \alpha$.

Let $G_n(l,x)=\sqrt{n}\{\hat{F}(l,x)-F_{\hat{\boldsymbol{\theta}}}(l,x)/c(\hat{\boldsymbol{\theta}})\}$ and $G_n^{(b)}(l,x)=n^{-1/2}\sum_j Z_j^{(b)}\hat{V}_j(l,x;\hat{\boldsymbol{\theta}})$

for $b=1,2,\ldots,B$.

Theorem 2: *Suppose that (R1) through (R8) listed in Emura & Konno (2012 CSDA) hold. Under* H_0, *and given B,*

$$(G_n, G_n^{(1)}, ..., G_n^{(B)}) \to (G_{\boldsymbol{\theta}}, G_{\boldsymbol{\theta}}^{(1)}, ..., G_{\boldsymbol{\theta}}^{(B)})$$

in $D\{(-\infty,\infty)^2\}^{\otimes(B+1)}$, *where* $G_{\boldsymbol{\theta}}$ *is the mean zero Gaussian process whose covariance for* $(l,x),(l^*,x^*)\in(-\infty,\infty)^2$ *is given as*

$$\begin{aligned}&Cov\{G_{\boldsymbol{\theta}}(l,x),G_{\boldsymbol{\theta}}(l^*,x^*)\}=E\{V_j(l,x;\boldsymbol{\theta})V_j(l^*,x^*;\boldsymbol{\theta})\}\\&=F_{\boldsymbol{\theta}}(l\wedge l^*,x\wedge x^*)/c(\boldsymbol{\theta})-F_{\boldsymbol{\theta}}(l,x)F_{\boldsymbol{\theta}}(l^*,x^*)/c(\boldsymbol{\theta})^2-\mathbf{g}_{\boldsymbol{\theta}}(l,x)'I^{-1}(\boldsymbol{\theta})\mathbf{g}_{\boldsymbol{\theta}}(l^*,x^*),\end{aligned}$$

where $a\wedge b\equiv\min(a,b)$, *and* $G_{\boldsymbol{\theta}}^{(1)},...,G_{\boldsymbol{\theta}}^{(B)}$ *are independent copies of* $G_{\boldsymbol{\theta}}$.

Simulation under the null

- Model for (L^O, X^O) : Bivariate normal

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_L \\ \mu_X \end{bmatrix} = \begin{bmatrix} 120\text{-}62.63 \\ 60.82 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_L^2 & \sigma_{LX} \\ \sigma_{LX} & \sigma_X^2 \end{bmatrix} = \begin{bmatrix} 19.64^2 & 19.64 \times 16.81 \rho_{LX} \\ 19.64 \times 16.81 \rho_{LX} & 16.81^2 \end{bmatrix}$$

- Generate data:

$$\{(L_j, X_j)\,(j = 1, \ldots, 400)\} \quad \text{subject to } L_j \le X_j$$

- Compute Cramér-von Mises statistic

$$C = \sum_j \{\, \hat{F}(L_j, X_j) - F_{\hat{\boldsymbol{\theta}}}(L_j, X_j) / c(\hat{\boldsymbol{\theta}}) \,\}^2$$

- Compute Bootstrap version

$$* \text{Multiplier Boot}: \quad \{\, C^{(b)}; b = 1, 2, \ldots, 1000 \,\}$$

$$* \text{Usual Parametric Boot}: \quad \{\, C^{*(b)}; b = 1, 2, \ldots, 1000 \,\}$$

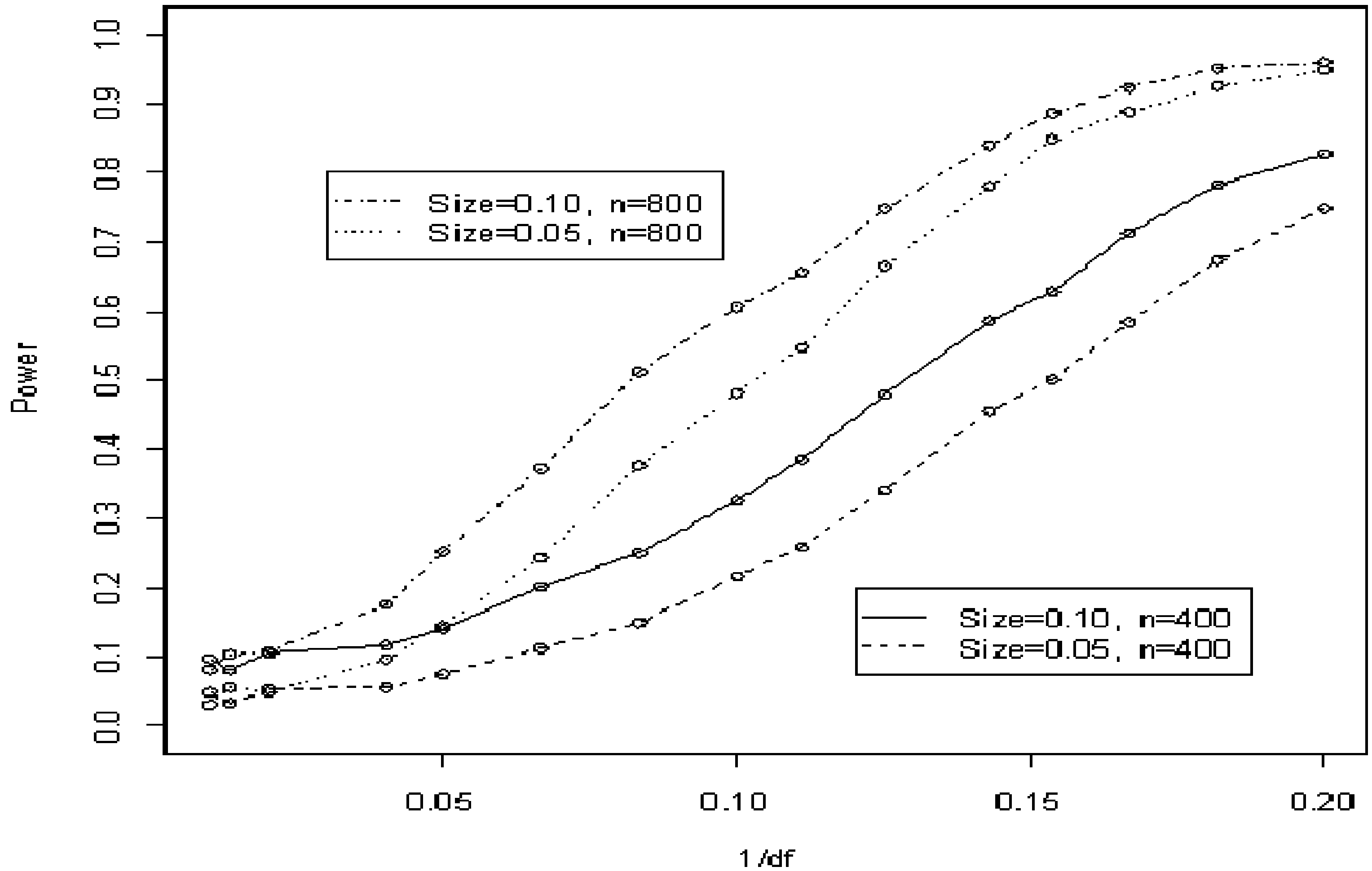
Multiplier Boot vs. naïve Bootstrap with $n = 400$

	Elapsed time (in second)		Resampling mean and SD (in parenthesis)	
	Multiplier	Parametric Bootstrap	Multiplier	Parametric Bootstrap
$\rho_{X^o Y^o}$ =0.70	23.17	1402.31	0.0621 (0.0314)	0.0603 (0.0304)
$\rho_{X^o Y^o}$ =0.35	22.84	1463.01	0.0606 (0.0269)	0.0581 (0.0252)
$\rho_{X^o Y^o}$ =0.00	23.04	1539.29	0.0597 (0.0219)	0.0586 (0.0225)
$\rho_{X^o Y^o}$ =-0.35	23.37	1380.09	0.0598 (0.0234)	0.0577 (0.0209)
$\rho_{X^o Y^o}$ =-0.70	26.77	1296.32	0.0591 (0.0225)	0.0569 (0.0217)

Almost same null distribution

Power study

we generated data from the bivariate *t*-distribution (Lang et al., 1989) while we performed the goodness-of-fit test under the null hypothesis of the bivariate normal distribution.



Data analysis

Law school data in US

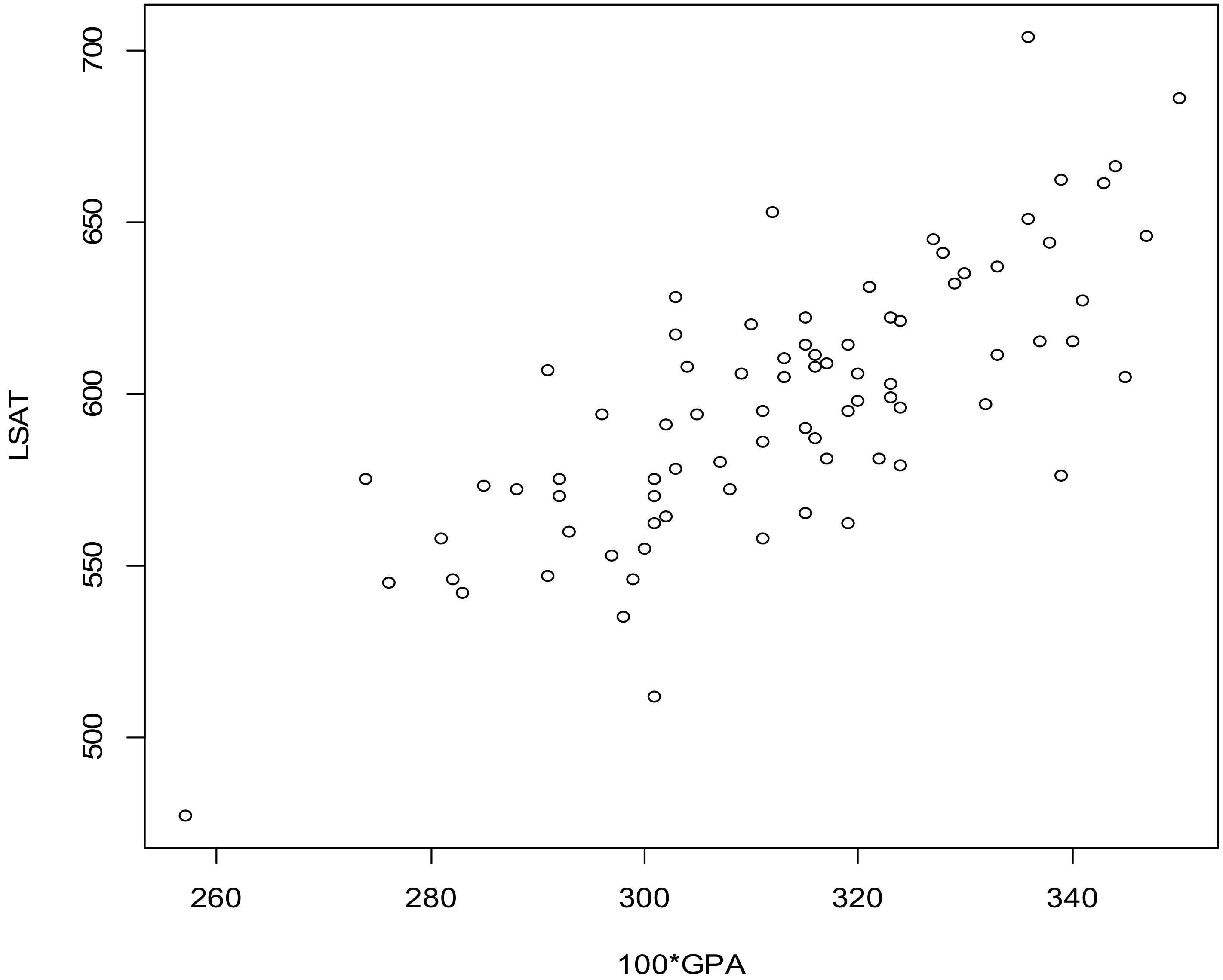
(Efron & Tibshirani 1993, *An Introduction to the Bootstrap*)

- Average LSAT (the national law test)
- Average GPA (graduate point average)

for N=82 American law schools

$$(\mathrm{LSAT}_j, \mathrm{GPA}_j) \quad \text{for} \quad j = 1, 2, \ldots, N$$

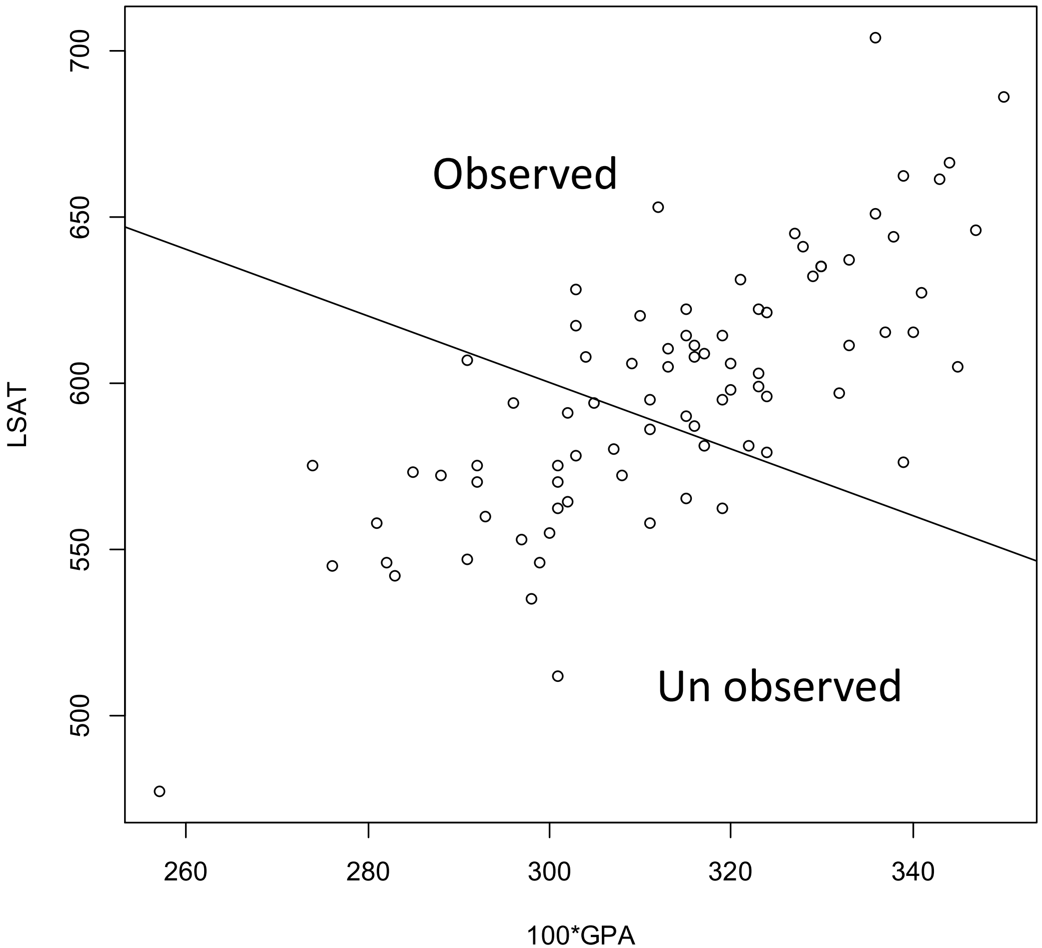
$$\{(\text{GPA}_j, \text{LSAT}_j); j = 1,\ldots,82\}$$



- We artificially define the truncation

(i) If $\text{LSAT} + 100 \times \text{GPA} >= 900$, we observe the sample

(ii) If $\text{LSAT} + 100 \times \text{GPA} < 900$, nothing observed!



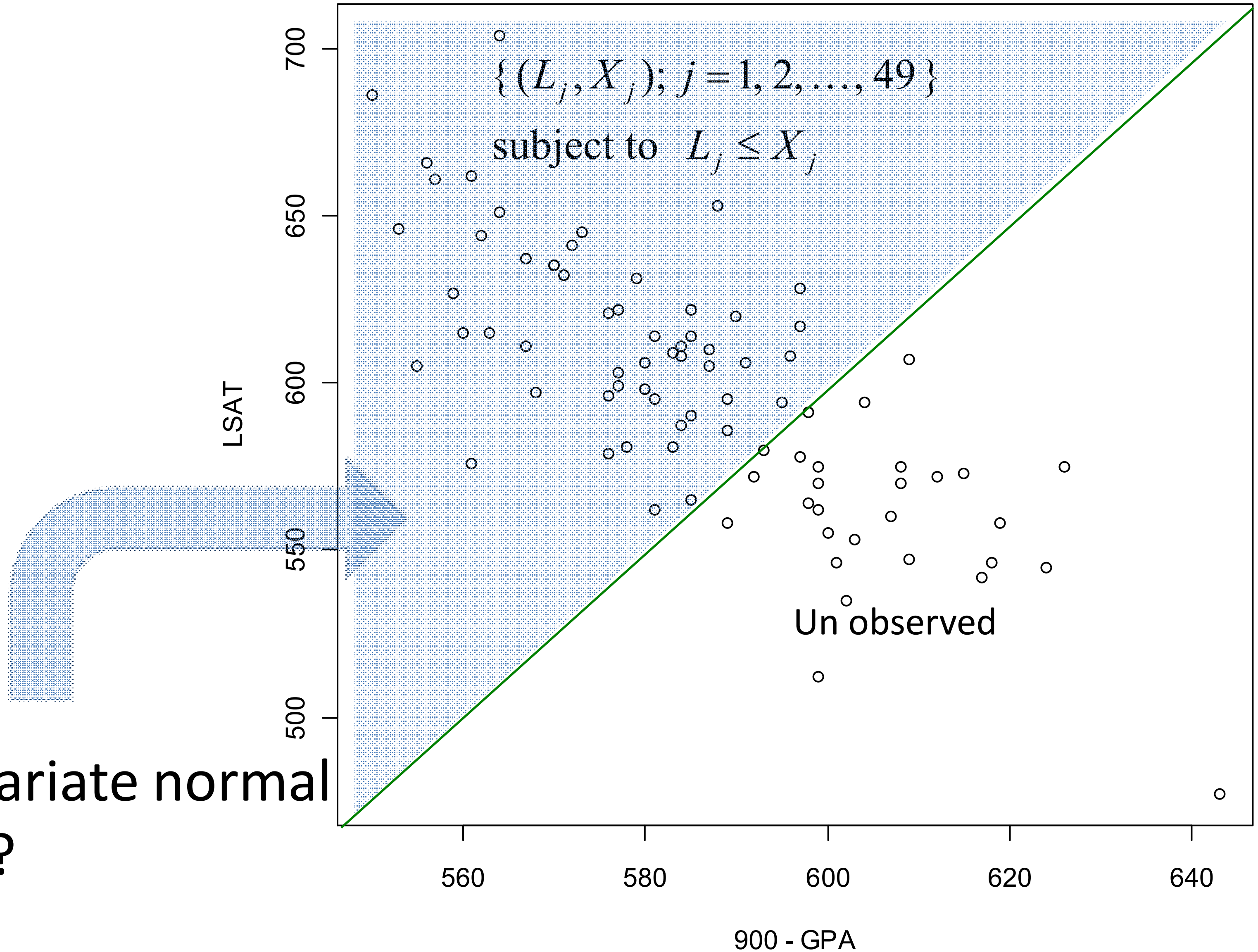
- We rewrite the truncation criterion

$$\begin{aligned} &\text{LSAT} + 100 \times \text{GPA} \ \geq \ 900, \\ &\Leftrightarrow \ \text{LSAT} \ \geq \ 900 \ - \ 100 \times \text{GPA} \\ &\Leftrightarrow \ X^O \ \geq \ L^O \end{aligned}$$

$$\text{where} \begin{cases} X^O = \text{LSAT} \\ L^O = \ 900 \ - \ 100 \times \text{GPA} \end{cases}$$

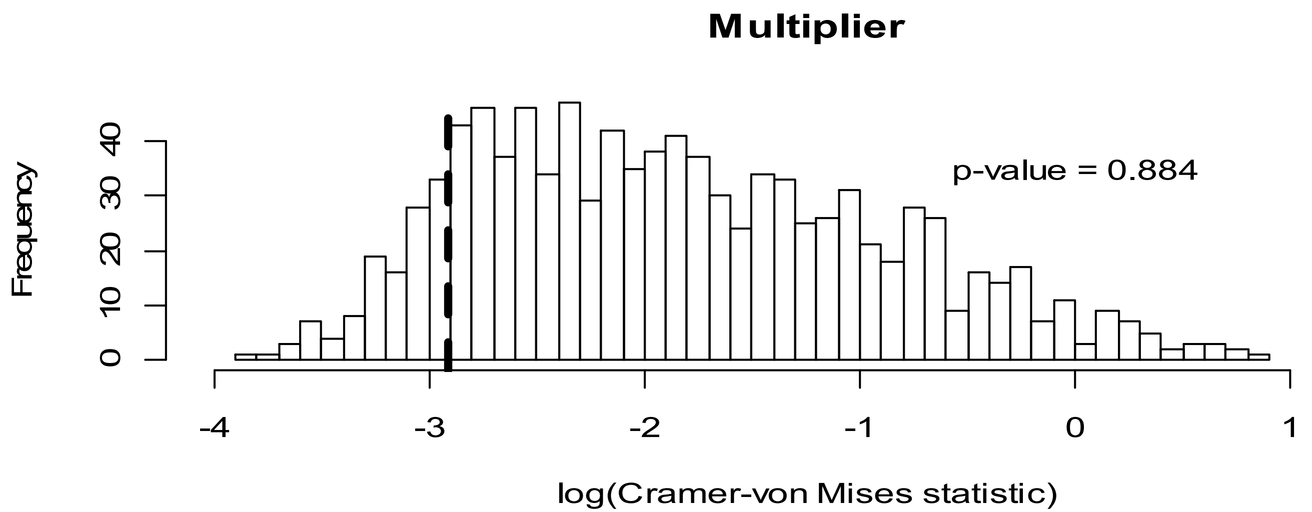
* Observe $\{ (L_j, X_j);\ j = 1, 2, \ldots, 49 \}$,

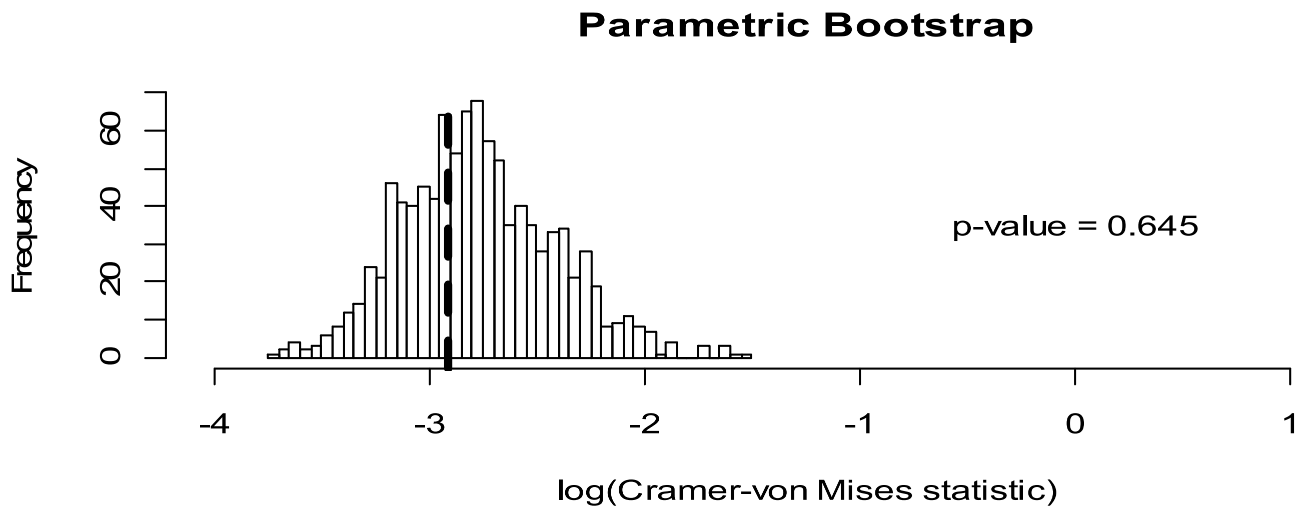
subject to $L_j \leq X_j$, where $L_j = 900 - 100 \times \text{GPA}_j$ and $X_j = \text{LSAT}_j$.



bivariate normal
fit ?

Computational time for the multiplier method = 1.25 second
Computational time for the parametric Bootstrap = 222.46 second





Summary: Why multiplier reduce computational cost ?

- The multiplier involves only arithmetic operations (multiply, sum). On the other hand, the parametric Bootstrap involves nonlinear maximization to get $\hat{\boldsymbol{\theta}}^{(b)}$
- Re-sampling from a truncated parametric model is involved:

Accept-reject algorithm

(i) data (L, X) from the distribution function $F_{\hat{\boldsymbol{\theta}}}(l, x)$ is generated;

(ii) if $L \leq X$, we accept the sample and set $(L_j^{(b)}, X_j^{(b)}) = (L, X)$;

otherwise we reject (L, X) and return to (i).

On the other hand, the multiplier only requires generating i.i.d. sequences

Thank you for your kind attention

- Bücher, A. and Dette, H. , 2010. A note on bootstrap approximations for the empirical copula process. Statist. Probab. Lett. 80, 1925-1932.
- Dietz, E., Böhning, D., 2000. On estimation of the Poisson parameter in zero-modified Poisson models. Computational Statistics & Data Analysis 34, 441-459.
- Efron, B., Tibshirani, R. J., 1993. An Introduction to the Bootstrap. Chapman & Hall, London.
- Emura, T., Konno, Y., 2012. A goodness-of-fit test for parametric models based on dependently truncated data, Computational Statistics & Data Analysis *56, 2237-2250*.
- Emura, T., Konno, Y, 2012. Multivariate normal distribution approaches for dependently truncated data. Statistical Papers *53 (No.1), 133-149.*
- Genest C., Remillard B., 2008. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, Annales de Institut Henri Poicare –Probabilites et Statistiques 44, 1096-1127.
- Henze N., 1996. Empirical-distribution-function goodness-of-fit tests for discrete models. Cand. J. Statist. 24, 81-93.
- Lilliefors HW, 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. JASA 62 399-402.
- Lilliefors HW, 1969. On the Kolmogorov-Smirnov test with exponential distribution with mean unknown. JASA 64 387-389.
- Fang, K.-T., Kotz, S., Ng, K. W., 1990. Symmetric Multivariate and Related Distributions. Chapman & Hall, New York.
- Holgate, 1964. Estimation of the bivariate Poisson distribution. Biometrika 51, 241-245.
- Klein, J. P., Moeschberger, M. L., 2003. Survival Analysis: Techniques for Censored and Truncated Data. New York, Springer.
- Kojadinovic, I., Yan, J., Holmes, M., 2011. Fast large-sample goodness-of-fit tests for copulas. Statistica Sinica 21, 841-871.
- Kojadinovic, I., Yan, J., 2011. A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. Statistics and Computing 21, 17-30.
- Lang, K. L., Little, J. A., Taylor, J. M. G., 1989. Robust statistical modeling using the t distribution. Journal of the American Statistical Association 84, 881-896.
- Stute W et al. , 1993. Bootstrap based goodness-of-fit test. Metrika 40, 243-256.
- van der Vaart. A. W., Wellner, J. A., 1996. Weak Convergence and Empirical Process, Springer Series in Statistics, Springer-Verlag, New York.
- J. P. Klein, M. L. Moeschberger, Survival Analysis: Techniques for Censored and Truncated Data, 2nd edition, Springer, New York, 2003.