# Bootstrap methods for structural relationship models

Juris Cielēns

University of Latvia

28.05.2011

# Location-scale model

## $H_0$ :

$$F_1(x) = F_2\left(\frac{x-\mu}{\sigma}\right), \ x \in \mathbb{R}$$

$$F_1^{-1}(t) = \sigma F_2^{-1}(t) + \mu, \ t \in [0,1]$$

## Transformation

$$\widetilde{Y}_i = \hat{\sigma} Y_i + \hat{\mu}$$

## $H_0$ for transformation

$$F_1(\widetilde{F}_2^{-1}(t)) = t, \ t \in [0,1]$$

# Location-scale model

**Density functions that belong to the location scale class**

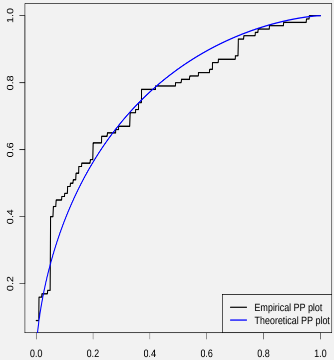
$$\text{Normal distribution: } f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Uniform distribution: } f(x,a,b) = \frac{1}{b-a} \text{ for } a \leq x \leq b$$

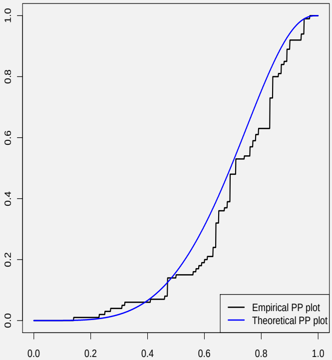
$$\text{Cauchy distribution: } f(x,\mu,\gamma) = \frac{1}{\pi}\left[\frac{\gamma}{(x-\mu)^2+\gamma^2}\right]$$

# Location-scale model

## PP plot example




N(0,1) against N(1,1) with sample sizes n=m=100




N(0,1) against N(-1,4) with sample sizes n=m=100

# Lehmann's alternatives model

**$H_0$:**

$$F_1(x) = 1 - (1 - F_2(x))^{1/h}, \ x \in \mathbb{R}$$

$$F_1^{-1}(t) = F_2^{-1}(1 - (1 - t)^h), \ t \in [0, 1]$$

**Transformation**

$$\widetilde{Y}_i = F_{2m}^{-1}(1 - (1 - F_{2m}(Y_i))^{\hat{h}})$$

**$H_0$ for transformation**

$$F_1(\widetilde{F}_2^{-1}(t)) = t, \ t \in [0, 1]$$

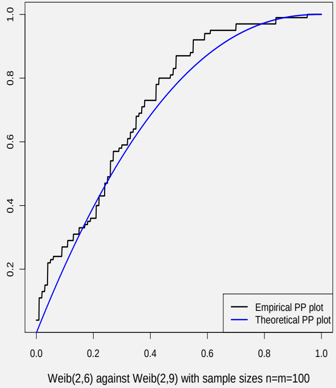
# Lehmann's alternatives model

**Density functions that belong to the Lehmann's alternative class**

Exponential distribution: $f(x, \lambda) = \lambda e^{-\lambda x}, \ x \geq 0$

Weibull distribution: $f(x, \lambda, k) = \frac{k}{\lambda} \left( \frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k}$

# PP plot example



# Parameter estimation using Mallows distance

## Location scale

$$M(F_1, F_2) := \int_0^1 (F_1^{-1}(t) - \sigma F_2^{-1}(t) - \mu)^2 dt$$

## Lehmann's alternatives

$$M(F_1, F_2) := \int_0^1 (F_1^{-1}(t) - F_2^{-1}(1 - (1 - t)^h)^2 dt$$

# Statistics

## Location scale

$$\sup_{0\le t\le 1} |\sqrt{n}(F_{1n}(\hat{\sigma}F_{2m}^{-1}(t)+\hat{\mu}) - F_1(\sigma F_2^{-1}(t)+\mu))|$$

or

$$\sup_{0\le t\le 1} |\sqrt{n}(F_{1n}(\widetilde{F}_{2m}^{-1}(t)) - F_1(\widetilde{F}_2^{-1}(t)))|$$

## Lehmann's alternatives

$$\sup_{0\le t\le 1} |\sqrt{n}(F_{1n}(F_{2m}^{-1}(1-(1-t)^{\hat{h}})) - F_1(F_2^{-1}(1-(1-t)^{h})))|$$

or

$$\sup_{0\le t\le 1} |\sqrt{n}(F_{1n}(\widetilde{F}_{2m}^{-1}(t)) - F_1(\widetilde{F}_2^{-1}(t)))|$$

# Statistics

**Assymptotical distribution for Location scale model**

$$\sup_{0\le t\le 1} |\sqrt{n}(F_{1n}(\hat{\sigma}F_{2m}^{-1}(t)+\hat{\mu}) - F_1(\sigma F_2^{-1}(t)+\mu))| \to$$

$$\sup_{0\le t\le 1}\left|B^{(n)}(t)+\sqrt{\frac{n}{m}}B^{(m)}(t)+\sqrt{n}f_2(F_2^{-1}(t))((\hat{\mu}-\mu)-F_2^{-1}(t)(\hat{\sigma}-\sigma))\right|$$

# First idea

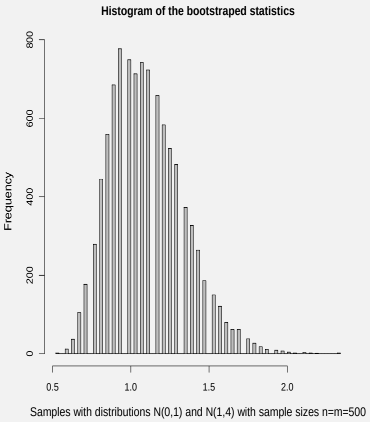
## Nonparametric Bootstrap

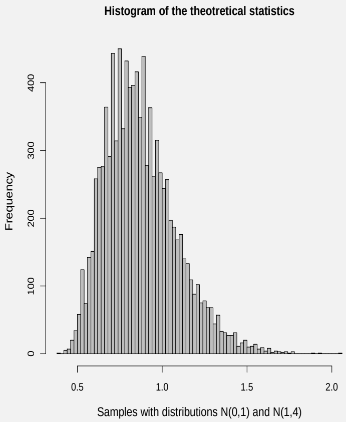
$$\sup_{0 \le t \le 1} |\sqrt{n}(F^*_{1n}(\hat{\sigma}^* {F^*_{2m}}^{-1}(t) + \hat{\mu}^*) - F_{1n}(\hat{\sigma} F^{-1}_{2m}(t) + \hat{\mu}))|$$

We bootstrap samples from $F_{1n}$ and $F_{2m}$.

# Bootstrap

## Result



# Smooth estimators

## Kernel definition

Function K is called a kernel if

- $\int_{-\infty}^{\infty} K(u)du = 1$;
- $\forall u\ K(-u) = K(u)$.

## Smoothed estimators

Density function: $\hat{f}_n(x) = 1/(nh) \sum_{i=1}^{n} K((x - X_i)/h)$;
Define $H(x) = \int_{-\infty}^{x} K(u)du$ then the smoothed distribution function is estimated as

$$\hat{F}_n(x) = 1/n \sum_{i=1}^{n} H((x - X_i)/h))$$

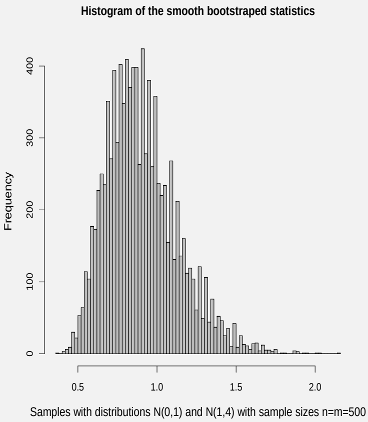
# Smooth bootstrap (L. Horwáth, 2008 )

$$\sup_{0\leq t\leq 1} |\sqrt{n}(\hat{F}^*_{1n}(\hat{\sigma}^*\hat{F}^{*}_{2m}{}^{-1}(t)+\hat{\mu}^*)-\hat{F}_{1n}(\hat{\sigma}\hat{F}^{-1}_{2m}(t)+\hat{\mu}))|$$

Samples are bootstrapped from $\hat{F}_{1n}$ and $\hat{F}_{2m}$.

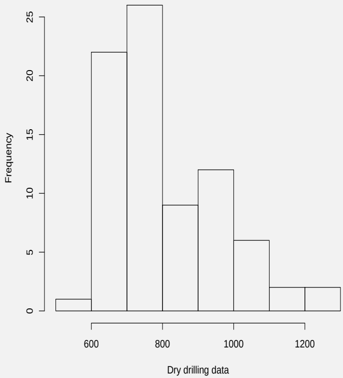
# Smooth bootstrap

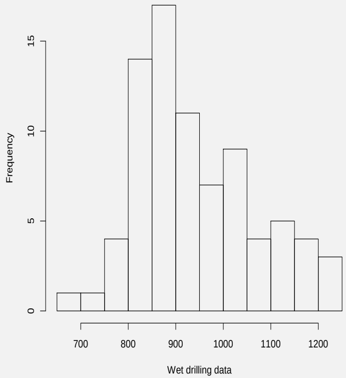
## Result



# Application

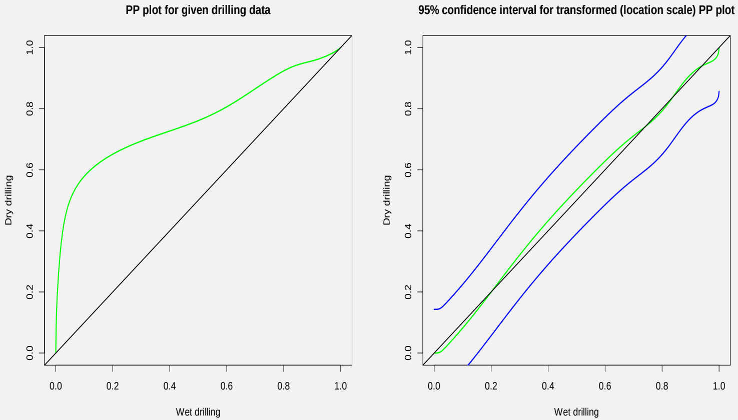
## Drilling data



# Application

## Drilling data



# Application

## Drilling data