

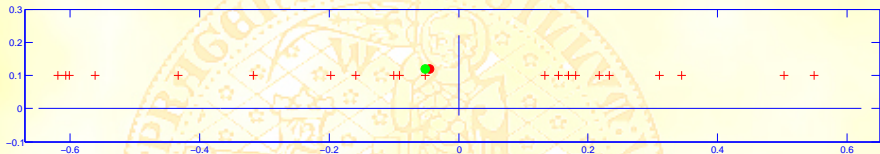


INSTITUTE OF ECONOMIC STUDIES, FACULTY OF SOCIAL SCIENCES  
CHARLES UNIVERSITY IN PRAGUE (*established 1348*)

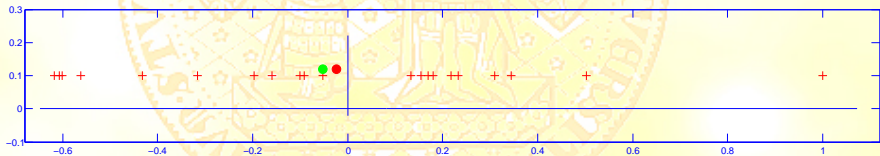
First of

Observe the mean • and the median •

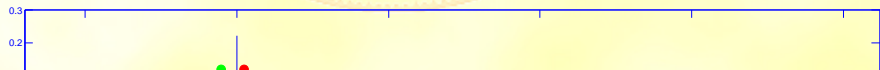
Non-contaminated data



at point 1



at point 2



Then there is of course a question:

Why we use more frequently mean than median?

$$\lim_{n \rightarrow \infty} \frac{\text{var}_{\Phi}(\bar{x}^{(n)})}{\text{var}_{\Phi}(\text{median}^{(n)})} = 0.6\dots$$

## Hampel's approach - characteristics of the functional $T$ at the d. f. $F$

- An **overall characteristic of the functional** (the estimator) is

$$\varepsilon^* = \sup \{ \varepsilon \leq 1 : \exists K_\varepsilon \subset \Theta, K_\varepsilon \text{ compact}$$

$$\left. \begin{array}{l} \pi(F, G) < \varepsilon \Rightarrow G(\{T_n \in K_\varepsilon\}) \xrightarrow{n \rightarrow \infty} 1 \end{array} \right\}$$

where  $\pi(F, G)$  is the *Prokhorov metric* of  $F(x)$  and  $G(x)$   
and  $T_n$  is an empirical counterpart to the functional  $T$ .

- $\varepsilon^*$  is called **breakdown point**

(explanation of Prokhorov metric is on one of the next slides,  
then finite sample breakdown point).

## Hampel's approach - characteristics of the functional $T$ at the d. f. $F$

- An overall characteristic of the functional (the estimator) is

 $\varepsilon^*$ 

It is not trivial to understand this definition -  
- we shall try after some preparation.

where  $\rho(T, G)$  is the Pitmanov metric of  $T(x)$  and  $G(x)$   
and  $T_n$  is an empirical counterpart to the functional  $T$ .

- $\varepsilon^*$  is called breakdown point.

## Hampel's approach - characteristics of the functional $T$ at the d. f. $F$

An overall characteristic of the functional (the estimator) is

The definition contains new notions:

compact set and Prokhorov metric.

- 1 Firstly, what is a compact set? And what is good for ?
- 2 Secondly, an inspiration why we need Prokhorov metric.
- 3 Thirdly, an explanation what Prokhorov metric is.

$\epsilon^*$  is called breakdown point.

## Hampel's approach - characteristics of the functional $T$ at the d. f. $F$

- An overall characteristic of the functional (the estimator) is

Then we return to the definition of **breakdown point**.

- 1 Firstly, we explain the sense of it.
- 2 Then we try to read the mathematics.

where  $\rho(T, G)$  is the *Prokhorov metric* of  $T(x)$  and  $G(x)$  and  $T_n$  is an empirical counterpart to the functional  $T$ .

- $\varepsilon^*$  is called breakdown point.

So, let us start with compact set

- Open and closed sets

$C$  is closed set if for any sequence  $\{x_n\}_{n=1}^{\infty} \subset C$  such that

$$\exists \left( \lim_{n \rightarrow \infty} x_n = x_0 \right) \Rightarrow x_0 \in C.$$

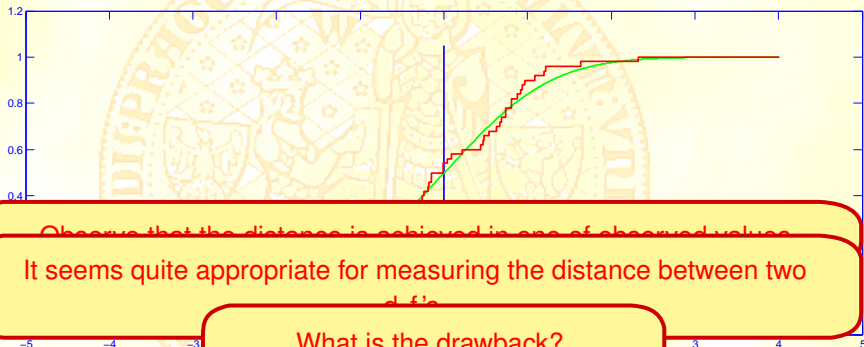
- $C$  is compact if it is closed and  $\exists K < \infty$   
so that  $\forall (x \in C) \|x\| < K < \infty$ .
- The sense of compact sets and the use of compactness:

Any open cover contains a finite subcover.



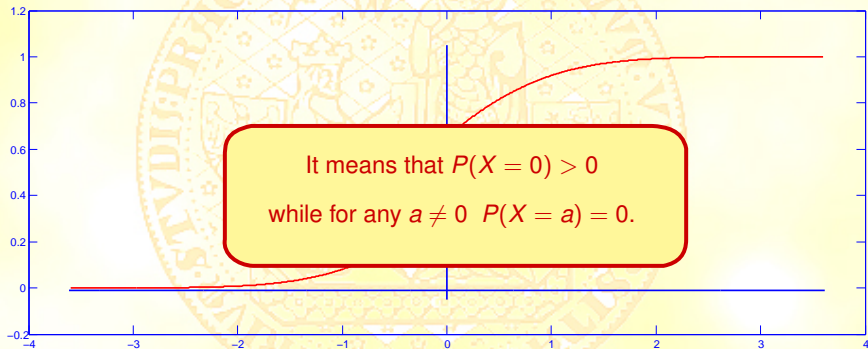
## Recalling Kolmogorov-Smirnov metric

$$D(F, G) = \sup_{-\infty < x < \infty} |F(x) - G(x)|$$



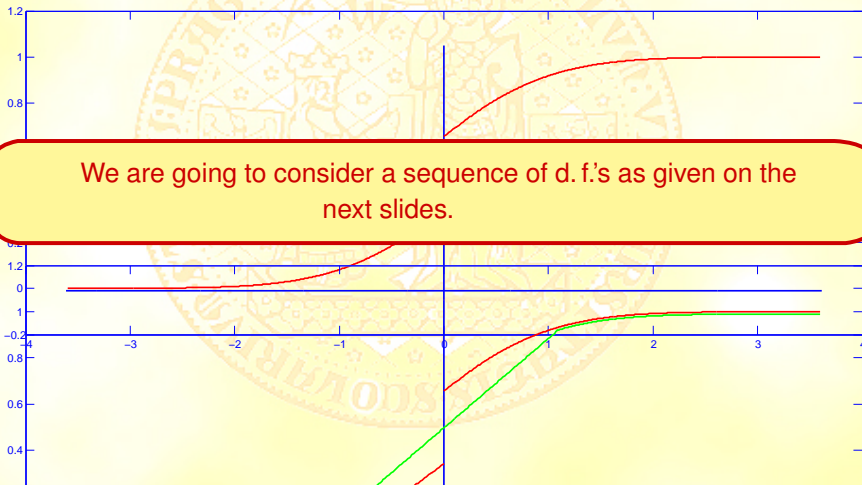
## Intuitive convergence of d. f.'s - but the Kolmogorov-Smirnov distance

Distribution function with a jump in the origine



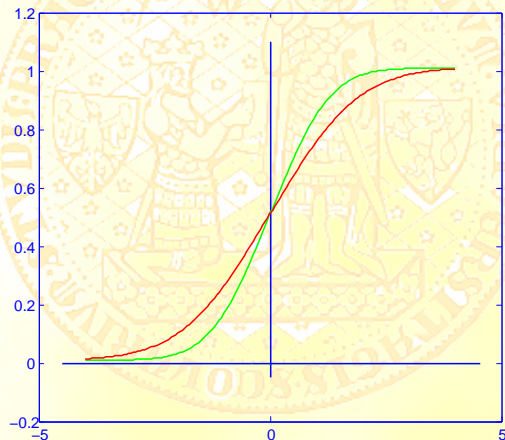
## Intuitive convergence of d. f.'s - but the Kolmogorov-Smirnov distance

D. f. with a jump in the origine

Width of linear part 2.21.8  
1.51.51.10.80.40.30.20.10.080.06.

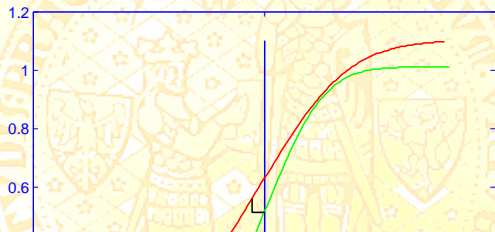
## Explaining Prohorov distance (notice different transcription)

Consider two d. f.'s

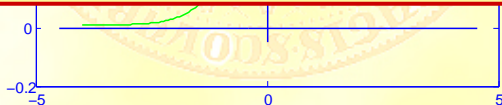


## Explaining Prokhorov distance

## Prokhorov metric



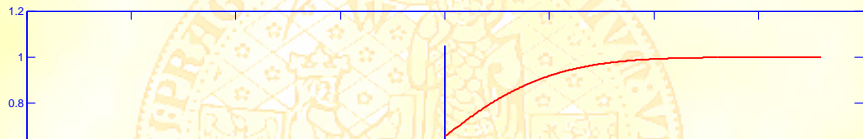
$$\pi(F, G) = \inf \{ \varepsilon \in [0, 1] : \forall (x \in \mathbb{R}) \quad F(x) \leq G(x + \varepsilon) + \varepsilon \vee G(x) \leq F(x + \varepsilon) + \varepsilon \}$$



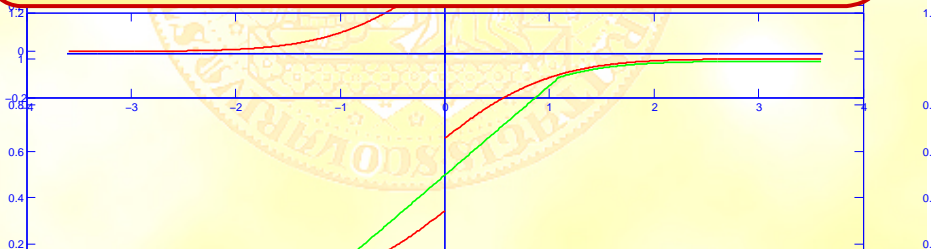
## Comparing Kolmogorov-Smirnov and Prokhorov metrics

D. f. with a jump in the origine Width of linear part 2.21.8

1.51.10.80.40.30.20.10.080.06.



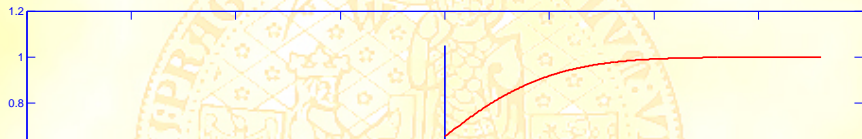
Let' recall the sequence we have considered.



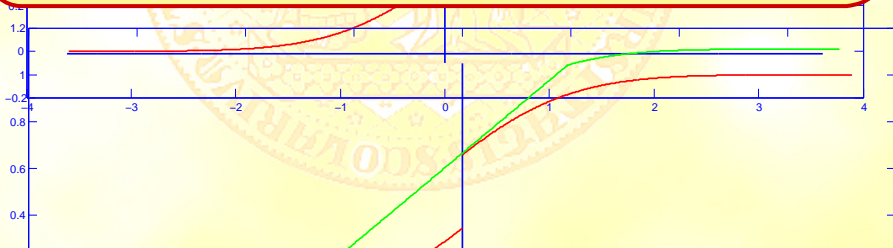
The convergence in Prokhorov metrics corresponds with intuitive idea of it.

D. f. with a jump in the origine

Width of linear part 2.21.8  
1.51.10.80.40.30.20.10.080.06.



Let's again consider a sequence of d. f.'s as given on the next slides - but now shifted to the left and up.



## The global empirical characteristic of estimator.

Hampel's approach - characteristics of the functional  $T$  at the d.f.  $F$

- Breakdown point - "finite" sample version

$$x_1, x_2, \dots, x_n \Rightarrow T_n(x_1, x_2, \dots, x_n)$$

- Find maximal  $m_n$  such that for any

$$y_1, y_2, \dots, y_{m_n} \Rightarrow |T_n(x_1, x_2, \dots, x_{n-m_n}, y_1, y_2, \dots, y_{m_n})| < \infty$$

$$(0 < T_n(x_1, x_2, \dots, x_{n-m_n}, y_1, y_2, \dots, y_{m_n}) < \infty \text{ - for scale}).$$

- Put

$$\varepsilon^* = \lim_{n \rightarrow \infty} \frac{m_n}{n}$$

(we'll return to it later on,

now let's return to the exact definition of the **breakdown point**).



## Let's read the definition of breakdown point

Hampel's approach - characteristics of the functional  $T$  at the d. f.  $F$ 

- An **overall characteristic of the functional** (the estimator) is

$$\varepsilon^* = \sup \{ \varepsilon \leq 1 : \exists K_\varepsilon \subset \Theta, K_\varepsilon \text{ compact} \}$$

$$\left. \pi(F, G) < \varepsilon \Rightarrow G(\{T_n \in K_\varepsilon\}) \xrightarrow{n \rightarrow \infty} 1 \right\}$$

where  $\pi(F, G)$  is the *Prokhorov metric* of  $F(x)$  and  $G(x)$   
and  $T_n$  is an empirical counterpart to the functional  $T$ .

- $\varepsilon^*$  is called **breakdown point**.

Let's rewrite the mathematical part of definition on the next slide.

## Let's read the definition of breakdown point

Hampel's approach - characteristics of the functional  $T$  at the d. f.  $F$ 

$$\varepsilon^* = \sup \{ \varepsilon \leq 1 : \exists K_\varepsilon \subset \Theta, K_\varepsilon \text{ compact}$$

$$\left. \pi(F, G) < \varepsilon \Rightarrow G(\{T_n \in K_\varepsilon\}) \xrightarrow{n \rightarrow \infty} 1 \right\}$$

- 1 Let's assume that  $T \in R$ , so it is connected with a parameter of  $F$ .
- 2  $K_\varepsilon$  could be sufficiently wide interval - for this moment.
- 3 The definition says that  $T_n$  is so "good" for estimating " $F$ " that whenever  $G$  is sufficiently close to  $F$ ,  $T_n$  converges in probability with respect to  $G$  to something finite.

That's all.

## Specifying the “robustness” characteristics for the location parameter

At the beginning of the third lecture

we have computed influence function for  $T(\Phi) = E_{\Phi}Z$ :

Recalling definition of influence function is :

Fix a functional  $T : \mathcal{H} \rightarrow R \dots$

$$IF(x, T, F) = \lim_{\delta \rightarrow 0} \frac{T\left((1 - \delta)F(\cdot) + \delta \cdot \Delta_x\right) - T\left(F(\cdot)\right)}{\delta}.$$


---

1 Fix  $T(\Phi) = E_{\Phi}(Z) = \int Z d\Phi = \int z \cdot \phi(z) dz$ .

2  $T(\Phi(\cdot)) = \frac{1}{\sqrt{2\pi}} \int z \cdot \exp\left\{-\frac{z^2}{2}\right\} dz = 0$

3  $T\left((1 - \delta)\Phi(\cdot) + \delta \cdot \Delta_x\right)$   
 $= \int z \left\{ (1 - \delta) \frac{1}{\sqrt{2\pi}} \cdot \exp\left\{-\frac{z^2}{2}\right\} + \delta \cdot \Delta_x \right\} dz = (1 - \delta) \cdot 0 + \delta \cdot x.$

4 Finally,  $IF(x, T, \Phi) = \lim_{\delta \rightarrow 0} \frac{\delta \cdot x}{\delta} = x.$

## Specifying the “robustness” characteristics for the location parameter

We easily verify that the same computation can be done  
whenever r. v.  $Z$  has finite mean value  $T(F) = E_F Z = \mu \in R$ :

Fix a functional  $T : \mathcal{H} \rightarrow R \dots$

$$IF(x, T, F) = \lim_{\delta \rightarrow 0} \frac{T\left((1 - \delta)F(\cdot) + \delta \cdot \Delta_x\right) - T\left(F(\cdot)\right)}{\delta}.$$


---

- 1 Fix  $T(F) = E_F(Z) = \int Z d\Phi = \int z \cdot f(z) dz$ .
- 2  $T(F(\cdot)) = \int z \cdot f(z) dz = \mu$
- 3  $T\left((1 - \delta)F(\cdot) + \delta \cdot \Delta_x\right)$   
 $= \int z \{(1 - \delta)f(x) + \delta \cdot \Delta_x\} dz = (1 - \delta) \cdot \mu + \delta \cdot x.$
- 4 Finally,  $IF(x, T, F) = \lim_{\delta \rightarrow 0} \frac{\delta \cdot (-\mu + x)}{\delta} = -\mu + x.$

## Specifying the “robustness” characteristics for the location parameter

Hence the “robustness” characteristics of  $T(F) = E_F(X)$  are:

- 1 The gross error sensitivity  $\gamma^* = \sup_{x \in R} |IF(x, T, F)| = \infty$ .
- 2 The local-shift sensitivity  $\lambda^* = \sup_{x, y \in R} \left| \frac{IF(x, T, F) - IF(y, T, F)}{x - y} \right| = 1$ .
- 3 The rejection point  $\rho^* = \inf \{r \in R^+ : IF(x, T, F) = 0, |x| > r\} = \infty$ .
- 4 The breakdown point  $\varepsilon^* = 0$

(the last characteristic is “derived heuristically”  
from the finite version of breakdown point).

## Specifying the “robustness” characteristics for the scale parameter

At the beginning of the third lecture

we have also computed influence function for  $T(\Phi) = E_{\Phi} X^2$ :

The **Recalling again definition of influence function is** :

Fix a functional  $T : \mathcal{H} \rightarrow R \dots$

$$IF(x, T, F) = \lim_{\delta \rightarrow 0} \frac{T\left((1 - \delta)F(\cdot) + \delta \cdot \Delta_x\right) - T\left(F(\cdot)\right)}{\delta}$$

- 1 Fix  $T(\Phi) = E_{\Phi}(X^2) = \int X^2 d\Phi = \int z^2 \cdot \phi(z) dz$ .
- 2  $T(\Phi(\cdot)) = \frac{1}{\sqrt{2\pi}} \int z^2 \cdot \exp\left\{-\frac{z^2}{2}\right\} dz = 1$
- 3  $T\left((1 - \delta)\Phi(\cdot) + \delta \cdot \Delta_x\right)$   
 $= \int z^2 \left\{ (1 - \delta) \frac{1}{\sqrt{2\pi}} \cdot \exp\left\{-\frac{z^2}{2}\right\} + \delta \cdot \Delta_x \right\} dz = (1 - \delta) \cdot 1 + \delta \cdot x^2$ .
- 4 Finally,  $IF(x, T, \Phi) = \lim_{\delta \rightarrow 0} \frac{(1 - \delta) \cdot 1 + \delta \cdot x^2 - 1}{\delta} = -1 + x^2$ .

## Specifying the “robustness” characteristics for the scale parameter

We easily verify that the same computation can be done whenever r.v.  $Z$  has finite variance  $T(F) = \mathbf{E}_F(Z - \mathbf{E}Z)^2 = \sigma^2 \in \mathbb{R}^+$ :

Fix a functional  $T : \mathcal{H} \rightarrow \mathbb{R} \dots$

$$IF(x, T, F) = \lim_{\delta \rightarrow 0} \frac{T\left((1 - \delta)F(\cdot) + \delta \cdot \Delta_x\right) - T(F(\cdot))}{\delta}.$$

- 1 Fix  $T(F) = \mathbf{E}_F(Z - \mathbf{E}Z)^2 = \int (Z - \mathbf{E}Z)^2 dF = \int (z - \mathbf{E}Z)^2 \cdot f(z) dz$ .
- 2  $T(F(\cdot)) = \int (z - \mathbf{E}Z)^2 \cdot f(z) dz = \sigma^2$
- 3  $T\left((1 - \delta)F(\cdot) + \delta \cdot \Delta_x\right) = \int (z - \mathbf{E}Z)^2 \{(1 - \delta)f(z) + \delta \cdot \Delta_x\} dz = (1 - \delta) \cdot \sigma^2 + \delta \cdot (x - \mathbf{E}Z)^2$ .
- 4 Finally,  $IF(x, T, F) = \lim_{\delta \rightarrow 0} \frac{\delta \cdot (-\sigma^2 + (x - \mathbf{E}Z)^2)}{\delta} = -\sigma^2 + (x - \mathbf{E}Z)^2$ .


## Specifying the “robustness” characteristics for the scale parameter

Hence the “robustness” characteristics of  $T(F) = E_F(Z - EZ)^2$  are:

- 1 The gross error sensitivity  $\gamma^* = \sup_{x \in R} |IF(x, T, F)| = \infty$ .
- 2 The local-shift sensitivity  $\lambda^* = \sup_{x, y \in R} \left| \frac{IF(x, T, F) - IF(y, T, F)}{x - y} \right| = \infty$ .
- 3 The rejection point  $\rho^* = \inf \{r \in R^+ : IF(x, T, F) = 0, |x| > r\} = \infty$ .
- 4 The breakdown point  $\varepsilon^* = 0$

(the last characteristic is again “derived heuristically”  
from the finite version of breakdown point).





*THANKS FOR ATTENTION*