Evidences of complexity of magnitude distribution, obtained from a non-parametric testing procedure

Stanislaw Lasocki

lasocki@geol.agh.edu.pl
Complex magnitude distributions

Singh, et al., 1983, BSSA 73, 1779-1796

Southern California

Knopoff, 2000, PNAS 97, 11880-11884

Main, 1995, BSSA 85, 1299-1308
How to reach statistical (measurable) significance of observational evidences?
How to reach statistical (measurable) significance of observational evidences?

Two problems with goodness of fit testing procedures:
How to reach statistical (measurable) significance of observational evidences?

Two problems with goodness of fit testing procedures:

1. \( H_0 (f_{act}(M) = f_{lin}(M)) \)
How to reach statistical (measurable) significance of observational evidences?

Two problems with goodness of fit testing procedures:

1. \( H_0 \left( f_{act}(M) = f_{lin}(M) \right) \)

   If \( \Pr(H_0) \) small \( \Rightarrow f_{act}(M) \neq f_{lin}(M) \) (likely)
How to reach statistical (measurable) significance of observational evidences?

Two problems with goodness of fit testing procedures:

1. \( H_0(f_{act}(M) = f_{lin}(M)) \)

   \( \text{If } \Pr(H_0) \text{ small} \Rightarrow f_{act}(M) \neq f_{lin}(M) \text{ (likely)} \)

   \( \text{but: } \sim \Rightarrow f_{act}(M) = f_{comp}(M) \)

   \( \text{because: } \sim \Rightarrow \exists f_{smooth}(M): f_{act}(M) = f_{smooth}(M) \)
How to reach statistical (measurable) significance of observational evidences?

Two problems with goodness of fit testing procedures:

1. \( H_0(f_{act}(M) = f_{lin}(M)) \)

   If \( \Pr(H_0)_{small} \Rightarrow f_{act}(M) \neq f_{lin}(M) \) (likely)

   but: \( \sim \Rightarrow f_{act}(M) = f_{comp}(M) \)

   because: \( \sim \Rightarrow \exists f_{smooth}(M): f_{act}(M) = f_{smooth}(M) \)

\( f_{lin}(M) \): exponential (G-R) \quad f_{smooth}^*(M) \): tapered G-R (Jackson, Kagan, 1999)

   gamma for \( M_0 \) (Kagan, 1999)

   generalized Pareto for \( M_0 \) (Pisarenko, Sornette, 2003)

   Weibull (Lasocki, 1993)

   double exponential (Lomnitz-Adler, Lomnitz, 1978)

   normal (Niazi, 1964)

   Utsu (1971), Makjanić (1972), Saito et al. (1973), Purcaru (1975), Seino et al. (1989) ...
2. Exponential-like shape of $f_{act}(M)$
2. Exponential-like shape of $f_{act}(M)$
2. Exponential-like shape of $f_{act}(M)$

???, • An artifact due to depopulation?
2. Exponential-like shape of $f_{act}(M)$

- An artifact due to depopulation?
- A statistical scatter?
2. Exponential-like shape of $f_{act}(M)$

- An artifact due to depopulation?
- A statistical scatter?
- A real break in scaling law?
Model-free testing: The smoothed bootstrap test for multimodality (Silverman, 1986; Efron, Tibshirani, 1998):
Model-free testing: The smoothed bootstrap test for multimodality (Silverman, 1986; Efron, Tibshirani, 1998):

\[ H_0^1(f_{act}(M) \text{ is unimodal}) \quad \text{and} \quad H_0^2(f_{act}(M) \text{ is unibumpal}) \]

Mode = a local maximum of PDF
Bump = \([a,b]\): PDF concave over \([a,b]\) and not over any larger interval
Model-free testing: The smoothed bootstrap test for multimodality (Silverman, 1986; Efron, Tibshirani, 1998):

$$H_0^1(f_{act}(M) \text{ is unimodal})$$

$$H_0^2(f_{act}(M) \text{ is unibumpal})$$

Mode = a local maximum of PDF

Bump = \([a,b]\): PDF concave over \([a,b]\) and not over any larger interval

Both, more than one mode and more than one bump in PDF are descriptive features indicating mixing of components. (Cox, 1966)

Low significance of either of $$H_0$$ evidences complexity of $$f_{act}(M)$$
TESTING PROCEDURE SCHEME

EQ DATA → RANDOMIZATION

TESTING FOR MULTIMODALITY
- Evaluating critical smoothing factor
- Smooth bootstrapping
- Estimating $H_0$ significance

TEST CALIBRATION
- Fitting G-R distribution model
- MC simulations
- Tests for multimodality of generated data sets

Calibrated significance of $H_0$
TESTING PROCEDURE SCHEME

EQ DATA

RANDOMIZATION

TESTING FOR MULTIMODALITY
- EVALUATING CRITICAL SMOOTHING FACTOR
- SMOOTH BOOTSTRAPPING
- ESTIMATING $H_0$ SIGNIFICANCE

TEST CALIBRATION
- FITTING G-R DISTRIBUTION MODEL
- MC SIMULATIONS
- TESTS FOR MULTIMODALITY OF GENERATED DATA SETS

CALIBRATED SIGNIFICANCE OF $H_0$
DATA RANDOMIZATION

To avoid a spurious complexity of $f_{act}(M)$ due to repetitions the observed magnitudes $M_{obs}$ are exponentially randomized within their round-off intervals $\delta M$ (=0.1)

$$ M_{obs} \overset{\text{\textcircled{R}}}{\rightarrow} M_{rand} $$

$$ M_{rand} = F_{e}^{-1} \left\{ u \left[ F_{e} (M_{obs} + 0.5\delta M) - F_{e} (M_{obs} - 0.5\delta M) \right] ight. $$

$$ + F_{e} (M_{obs} - 0.5\delta M) \right\} $$

$$ F_{e} (M) = 1 - \exp\left[- \beta (M - M_{c}) \right] $$

$u : Unif([0,1])$
Testing Procedure Scheme

**EQ DATA**

**Randomization**

**Testing for Multimodality**
- Evaluating Critical Smoothing Factor
- Smooth Bootstrapping
- Estimating $H_0$ Significance

**Test Calibration**
- Fitting G-R Distribution Model
- MC Simulations
- Tests for Multimodality of Generated Data Sets

**Calibrated Significance of $H_0$**
TESTING PROCEDURE SCHEME

**EQ DATA** → **RANDOMIZATION**

**TESTING FOR MULTIMODALITY**
- Evaluating Critical Smoothing Factor
- Smooth Bootstrapping
- Estimating $H_0$ Significance

**TEST CALIBRATION**
- Fitting G-R Distribution Model
- MC Simulations
- Tests for Multimodality of Generated Data Sets

**CALIBRATED SIGNIFICANCE OF $H_0$**
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor
Nonparametric, kernel $PDF$ estimate:

$$f_n(M | \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{M - M_i}{h}\right)$$

$$K(\xi) = (2\pi)^{-0.5} \exp\left(-\frac{\xi^2}{2}\right)$$

$M_i, i = 1,..,n$ – data
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

\[ f_n(M \mid \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{M - M_i}{h}\right) \]

\[ K(\xi) = (2\pi)^{-0.5} \exp\left(-\frac{\xi^2}{2}\right) \quad \text{kernel} \]

\[ M_i, i = 1, \ldots, n \quad \text{data} \]
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

\[ f_n(M \mid \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{M - M_i}{h}\right) \]

\[ K(\xi) = (2\pi)^{-0.5} \exp\left(-\frac{\xi^2}{2}\right) \quad \text{kernel} \]

\[ M_i, i = 1,\ldots,n \quad \text{data} \]

Its shape depends on \( h \):
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

\[ \hat{f}_n(M \mid \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K\left( \frac{M - M_i}{h} \right) \]

\[ K(\xi) = (2\pi)^{-0.5} \exp\left( -\frac{\xi^2}{2} \right) \]

\[ M_i, i = 1, \ldots, n \quad \text{– data} \]
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

$$f_n (M \mid \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K \left( \frac{M - M_i}{h} \right)$$

$$K(\xi) = (2\pi)^{-0.5} \exp \left( -\frac{\xi^2}{2} \right)$$ — kernel

$$M_i, i = 1, \ldots, n$$ — data

Its shape depends on \( h \):
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

\[ \hat{f}_n(M \mid \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K \left( \frac{M - M_i}{h} \right) \]

\[ K(\xi) = (2\pi)^{-0.5} \exp \left( -\frac{\xi^2}{2} \right) \]

Its shape depends on \( h \):

Data point for \( h = 0.3 \).
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

\[ f_n(M | \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K\left( \frac{M - M_i}{h} \right) \]

\[ K(\xi) = (2\pi)^{-0.5} \exp\left( -\frac{\xi^2}{2} \right) \quad \text{kernel} \]

\[ M_i, i = 1,..,n \quad \text{data} \]

Its shape depends on \( h \):
TESTING FOR MULTIMODALITY
Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

\[ \hat{f}_n(M | \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{M - M_i}{h}\right) \]

\[ K(\xi) = (2\pi)^{-0.5} \exp\left(-\frac{\xi^2}{2}\right) \]

\( M_i, i = 1, \ldots, n \) — data

Its shape depends on \( h \):
TESTING FOR MULTIMODALITY

Evaluating critical smoothing factor

Nonparametric, kernel PDF estimate:

\[ \hat{f}_n(M \mid \{M_i\}, h) = \frac{1}{nh} \sum_{i=1}^{n} K\left( \frac{M - M_i}{h} \right) \]

\[ K(\xi) = (2\pi)^{-0.5} \exp\left(-\frac{\xi^2}{2}\right) \quad \text{kernel} \]

\[ M_i, i = 1, \ldots, n \quad \text{data} \]

Its shape depends on \( h \):

Critical smoothing factor \( h_{cr} \):

\[ \hat{f}_n(M \mid \{M_i\}, h) \to \begin{cases} \text{unimodal for } h \geq h_{cr} \\ \text{multimodal for } h < h_{cr} \end{cases} \]
TESTING FOR MULTIMODALITY
Smooth bootstrapping
TESTING FOR MULTIMODALITY

Smooth bootstrapping

Smooth bootstrapping = Sampling from $\hat{f}_n (M \mid \{M_i\}, h_{cr})$
TESTING FOR MULTIMODALITY
Smooth bootstrapping

Smooth bootstrapping = Sampling from $\hat{f}_n(M \mid \{M_i\}, h_{cr})$

Standard bootstrap: $\{M_i\}_{i=1,\ldots,n} \overset{\text{®}}{\rightarrow} \{M'_i\}_{i=1,\ldots,n}$
TESTING FOR MULTIMODALITY
Smooth bootstrapping

Smooth bootstrapping = Sampling from $\hat{f}_n(M \mid \{M_i\}, h_{cr})$

Standard bootstrap: $\{M_i\}_{i=1,...,n} \overset{\circ}{\circ} \{M'_i\}_{i=1,...,n}$

Smooth bootstrap:

$\{M_i^{Silv}\}_{i=1,...,n} : M_i^{Silv} = M_i' + h_{cr} \varepsilon_i$

Silverman (1986):

$\{M_i^{Efr}\}_{i=1,...,n} : M_i^{Efr} = M_i^{Silv} + \frac{M_i^{Silv} - M_i^{Silv}}{\sqrt{1 + h_{cr}^2 / \sigma^2}}$

Efron, Tibshirani (1998):

$\varepsilon : \text{Norm}(0,1) \quad \sigma^2 = \text{samplevar}(M_i)$
TESTING FOR MULTIMODALITY

Estimating $H_0$ significance

\[ \hat{P}_{Silv}(H_0) = \frac{\text{the number of unimodal } \hat{f}_n(M \mid \{M_i^{Silv}\} h_{cr})}{R} \]

\[ \hat{P}_{Efr}(H_0) = \frac{\text{the number of unimodal } \hat{f}_n(M \mid \{M_i^{Efr}\} h_{cr})}{R} \]
TESTING PROCEDURE SCHEME

EQ DATA → RANDOMIZATION

TESTING FOR MULTIMODALITY
- Evaluating critical smoothing factor
- Smooth bootstrapping
- Estimating $H_0$ significance

TEST CALIBRATION
- Fitting G-R distribution model
- MC simulations
- Tests for multimodality of generated data sets

Calibrated significance of $H_0$
TESTING PROCEDURE SCHEME

EQ DATA → RANDOMIZATION

TESTING FOR MULTIMODALITY
- Evaluating critical smoothing factor
- Smooth bootstrapping
- Estimating $H_0$ significance

TEST CALIBRATION
- Fitting G-R distribution model
- MC simulations
- Tests for multimodality of generated data sets

Calibrated significance of $H_0$
TEST CALIBRATION
TEST CALIBRATION

Test result \( \hat{P}(H_0) = p^* \)
TEST CALIBRATION

Test result \( \hat{P}(H_0) = p^* \)

If \( H_0 \) were true how many samples like \( \{M_i\} \) would result in \( \hat{P}(H_0) \leq p^* \)?
TEST CALIBRATION
Fitting G-R distribution model
TEST CALIBRATION

Fitting G-R distribution model

\[ \{M_i\} \Rightarrow \hat{F}_{GR}(M) = \begin{cases} 
0 & M < M_c \\
\frac{1 - \exp\left[-\hat{\beta} \left(M - M_c\right)\right]}{1 - \exp\left[-\hat{\beta} \left(M_{\text{max}} - M_c\right)\right]} & M_c \leq M \leq M_{\text{max}} \\
1 & M > M_{\text{max}} 
\end{cases} \]

\[ \hat{M}_{\text{max}} = M_{\text{max}}^{\text{obs}} + \int_{M_c}^{M_{\text{max}}} [\hat{F}_{GR}(M)]^n dM \]

\[ M_{\text{max}}^{\text{obs}} = \max_i (M_i) \]
TEST CALIBRATION
Fitting G-R distribution model

\[ \{M_i\} \Rightarrow \hat{F}_{GR}(M) = \begin{cases} 
0 & \text{if } M < M_c \\
1 - \exp\left[-\frac{\beta (M - M_c)}{\hat{M}_{max} - M_c}\right] & \text{if } M_c \leq M \leq \hat{M}_{max} \\
1 & \text{if } M > \hat{M}_{max}
\end{cases} \]

\[ \hat{M}_{max} = M_{max}^{obs} + \int_{M_c}^{\hat{M}_{max}} [\hat{F}_{GR}(M)]^n dM \]

\[ M_{max}^{obs} = \max_i (M_i) \]

\[ \hat{F}_{GR}(M): \text{ - supports } H_0 \]

- could underlay \( \{M_i\} \)
TEST CALIBRATION
\[ \hat{F}_{GR}(M)^\circ \left\{ M_i^{GR} \right\}_{i=1,\ldots,n} \downarrow_{k=1,\ldots,L} \]
TEST CALIBRATION

\[
\hat{F}_{GR}(M) \subseteq \bigcup_{i=1}^{n} \bigcap_{k=1}^{L} \{M_{i}^{GR}\} \\
\forall \{M_{i}^{GR}\}_{i=1}^{n} \Rightarrow H_{0} \text{ test } \Rightarrow L \text{ values of } \hat{P}(H_{0})
\]
TEST CALIBRATION

\[ \hat{F}_{GR} (M) \overset{\text{®}}{\supseteq} \{ M_i^{GR} \} = 1, \ldots, n \]

\[ \forall \{ M_i^{GR} \}_{i=1}^{n} \overset{\text{®}}{\supseteq} H_0 \text{ test} \Rightarrow L \text{ values of } \hat{P}(H_0) \]

\[ L \text{ values } \hat{P}(H_0) \overset{\text{®}}{\supseteq} \forall p^* \ n = \frac{\text{No of } \hat{P}(H_0) < p^*}{L} \]
TEST CALIBRATION

\[ \hat{F}_{GR}(M) \overset{\text{REG}}{\left\{M_i^{GR}\right\}}_{i=1,\ldots,n} \mathcal{L} \]

\[ \forall \left\{M_i^{GR}\right\}_{i=1,\ldots,n} \overset{\text{REG}}{\mathcal{L}} H_0 \text{ test } \Rightarrow L \text{ values of } \hat{P}(H_0) \]

\[ L \text{ values } \hat{P}(H_0) \overset{\text{REG}}{\forall p^* \ \nu = \frac{\text{No of } \hat{P}(H_0) < p^*}{L}} \]

\[ p^* \text{ vs } \nu : \text{ calibration curve} \]
SELECTED RESULTS

• REGIONAL CATALOGS

Southern-California earthquakes (*SCSN Format EQ Catalog*)

983 EQ-s from 1.07.1944-1.03.1990 $M \geq 4.0$

Area definition: Nordquist (1964)
Sample selection: Knopoff (2000)

\[
\text{Prob}\{H_0^1: \text{unimodality}\} = 0.102 \\
\text{Prob}\{H_0^2: \text{one bump}\} = 0.091
\]
SELECTED RESULTS

• REGIONAL CATALOGS

Northern-California earthquakes: Northern California Earthquake Catalog and Phase Data (Northern California Seismic Network, U.S. Geological Survey, Menlo Park; Berkeley Seismological Laboratory, University of California, Berkeley)

603 EQ-s from 1.01.1968-31.12.2006 $M \geq 4.0$

lat$>38^\circ$

Prob\{$H^1_0$: unimodality\} = 0.042
Prob\{$H^2_0$: one bump\} = 0.074
SELECTED RESULTS

- REGIONAL CATALOGS - GREECE
SELECTED RESULTS

• REGIONAL CATALOGS - GREECE

Central Ionian Islands earthquakes:
1256 EQ-s from 1981-2001 \( M \geq 4.0 \)

\[
\text{Prob}\{H_0^1: \text{unimodality}\} = 0.069
\]
\[
\text{Prob}\{H_0^2: \text{one bump}\} = 0.080
\]
SELECTED RESULTS

- REGIONAL CATALOGS - GREECE

Central Ionian Islands earthquakes:
1256 EQ-s from 1981-2001 $M \geq 4.0$

\[
\text{Prob}\{H_0^1: \text{unimodality}\} = 0.069 \\
\text{Prob}\{H_0^2: \text{one bump}\} = 0.080
\]

Northern Aegean earthquakes:
744 EQ-s from 1981-2001 $M \geq 4.0$

\[
\text{Prob}\{H_0^1: \text{unimodality}\} = 0.049 \\
\text{Prob}\{H_0^2: \text{one bump}\} = 0.076
\]
SELECTED RESULTS

- REGIONAL CATALOGS - GREECE

Central Ionian Islands earthquakes:
1256 EQ-s from 1981-2001 $M \geq 4.0$

$\text{Prob}\{H_0^1: \text{unimodality}\} = 0.069$
$\text{Prob}\{H_0^2: \text{one bump}\} = 0.080$

Northern Aegean earthquakes:
744 EQ-s from 1981-2001 $M \geq 4.0$

$\text{Prob}\{H_0^1: \text{unimodality}\} = 0.049$
$\text{Prob}\{H_0^2: \text{one bump}\} = 0.076$

Thessalia earthquakes:
104 EQ-s from 1981-2001 $M \geq 4.0$

$\text{Prob}\{H_0^1: \text{unimodality}\} = 0.51$
$\text{Prob}\{H_0^2: \text{one bump}\} = 0.23$
SELECTED RESULTS

- REGIONAL CATALOGS – GREECE

Central Ionian Islands earthquakes with aftershocks removed (Reasenberg, 1985):

595 EQ-s from 1981-2001 $M \geq 4.0$

$\text{Prob}\{H_0^1: \text{unimodality}\} = 0.050$
$\text{Prob}\{H_0^2: \text{one bump}\} = 0.062$
SELECTED RESULTS

- WORLDWIDE CATALOGS

Large ($M \geq 7.0$), shallow ($h \leq 70\text{km}$), worldwide earthquakes: combined Pacheco-Sykes catalog (698 EQ-s, *Pacheco and Sykes, 1992*) and Harvard CMTS catalogs:

821 EQ-s from 1900-2002

\[
\text{Prob}\{H_0^1: \text{unimodality}\} = 0.362
\]

\[
\text{Prob}\{H_0^2: \text{one bump}\} = 0.096
\]
SELECTED RESULTS

- WORLDWIDE CATALOGS

Worldwide earthquakes: Harvard CMTS catalog

1825 EQ-s from 1.01.1977-31.12.2004 $M_w \geq 6.0$

$\text{Prob}\{H_0^1: \text{unimodality}\} = 0.162$

$\text{Prob}\{H_0^2: \text{one bump}\} < 9 \times 10^{-4}$
CONCLUSIONS
CONCLUSIONS

• The smoothed bootstrap test for multimodality makes it possible to investigate the complexity of earthquake magnitude distribution without making any specific assumptions on the distribution model.
CONCLUSIONS

• The smoothed bootstrap test for multimodality makes it possible to investigate the complexity of earthquake magnitude distribution without making any specific assumptions on the distribution model.

• The test evidence that the earthquake magnitude distribution frequently neither follows the Gutenberg-Richter law nor is smoothly non-linear but it is complex. Regarding magnitudes, earthquake populations are often not homogeneous. Traces of the complexity are more distinct for the regional than the worldwide data.
CONCLUSIONS

• The smoothed bootstrap test for multimodality makes it possible to investigate the complexity of earthquake magnitude distribution without making any specific assumptions on the distribution model.

• The test evidence that the earthquake magnitude distribution frequently neither follows the Gutenberg-Richter law nor is smoothly non-linear but it is complex. Regarding magnitudes, earthquake populations are often not homogeneous. Traces of the complexity are more distinct for the regional than the worldwide data.

• The complexity of magnitude distribution has also important implications for probabilistic seismic hazard assessment. When the actual magnitude distribution is complex and nonlinear features occur in a large magnitude region, the use of the presently known magnitude distribution models may yield unacceptable inaccuracy of the hazard estimates.
CENTRAL IONIAN ISLANDS

return period [years]

magnitude

Nonparametric
Gutenberg-Richter
REFERENCES

REFERENCES


