Testing for Poisson Behavior

Philip B. Stark
Department of Statistics, UC Berkeley
joint with Brad Luen

17 April 2012
Seismological Society of America
Annual Meeting
San Diego, CA

Quake Physics versus Quake Statistics

- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)
- Foreshocks, aftershocks, swarms—no physics-based definitions
- Spatially inhomogeneous temporally homogeneous Poisson model (SITHP) doesn't fit at regional scales
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- More complex models "motivated by physics."

Declustering and Poisson Behavior

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: “Ok, so why do you decluster the catalog?”

A: “to get the best possible estimate for the rate of mainshocks”

“the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence.”

- What’s a mainshock? Why does the distinction matter?
- Are declustered catalogs consistent with a Poisson model? With (conditional) independence of times and locations?

Declustering and Poisson Behavior

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: "Ok, so why do you decluster the catalog?"

A: "to get the best possible estimate for the rate of mainshocks"

"the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence."

- What's a mainshock? Why does the distinction matter?
- Are declustered catalogs consistent with a Poisson model? With (conditional) independence of times and locations?

Declustering and Poisson Behavior

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: "Ok, so why do you decluster the catalog?"

A: "to get the best possible estimate for the rate of mainshocks"

"the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence."

- What's a mainshock? Why does the distinction matter?
- Are declustered catalogs consistent with a Poisson model? With (conditional) independence of times and locations?

Declustering and Poisson Behavior

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: "Ok, so why do you decluster the catalog?"

A: "to get the best possible estimate for the rate of mainshocks"

"the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence."

- What's a mainshock? Why does the distinction matter?
- Are declustered catalogs consistent with a Poisson model? With (conditional) independence of times and locations?

“Main events,” “foreshocks,” and “aftershocks”

- An event that the declustering method does not remove is a main shock.
- An event that the declustering method removes is a foreshock or an aftershock.

. . . profound shrug . . .

Where’s the physics?

Focus here on statistics: Testing for Poisson behavior and for conditional independence of locations and times.

Poisson dispersion test (PD)

- Divide time into N_w intervals.
- Times conditionally IID, so N events are independent "trials" with N_w possible outcomes.
- Chance event falls in each interval is equal.
- Joint distribution of counts in intervals multinomial.
- Count of events in interval k is N_k.
- Expected number in each interval is $\bar{N} \equiv n/N_w$.
- PD chi-square statistic proportional to sample variance of counts:
  $\chi^2_{PD} \propto \sum_{k=1}^{N_w} (N_k - \bar{N})^2$.
- Classical chi-square goodness of fit test.
- Null distribution approximately chi-square; can estimate P-value more accurately by simulation.

Brown-Zhao test (BZ)

Same as PD test, but use a transformation of the counts:

$$Y_k \equiv \sqrt{N_k + 3/8}, \quad \bar{Y} \equiv \sum Y_k / N_w.$$

$$\chi^2_{BZ} \equiv 4 \sum_{k=1}^{N_w} (Y_k - \bar{Y})^2.$$

Under SITHP hypothesis, χ^2_{BZ} has a distribution that is approximately chi-square with $N_w - 1$ degrees of freedom.
Chi-square approximation to the null distribution of χ^2_{BZ} tends to be better than the chi-square approximation to the null distribution of χ^2_{PD}.

However, can estimate P-values for either by simulation instead.

Multinomial chi-square test (MC)

- Divide time into N_w intervals.
- In each interval, count of events unconditionally Poisson.
- Estimate rate λ of Poisson from observed total but pretend rate known a priori

$$K^- \equiv \min \left\{ k : \; N_w e^{-\lambda} \sum_{j=0}^{k} \lambda^j / j! \geq 5 \right\}.$$

$$K^+ \equiv \max \left\{ k : \; N_w \left(1 - e^{-\lambda} \sum_{j=0}^{k-1} \lambda^j / j! \right) \geq 5 \right\}.$$

Multinomial chi-square test, continued

Define

$$E_k \equiv \begin{cases} N_w e^{-\lambda} \sum_{j=0}^{K^-} \lambda^j/j!, & k = K^- \\ N_w e^{-\lambda} \lambda^k/k!, & k = K^- + 1, \ldots, K^+ - 1 \\ N_w (1 - e^{-\lambda} \sum_{j=0}^{K^+ - 1} \lambda^j/j!), & k = K^+. \end{cases}$$

$$X_k \equiv \begin{cases} \text{\# intervals with } \le K^- \text{ events}, & k = K^- \\ \text{\# intervals with } k \text{ events}, & k = K^- + 1, \ldots, K^+ - 1 \\ \text{\# intervals with } \ge K^+ \text{ events}, & k = K^+. \end{cases}$$

Test statistic: $\chi^2_{MC} \equiv \sum_{k=K^-}^{K^+} (X_k - E_k)^2/E_k.$

Not classical chi-square goodness of fit: “expected” estimated from data; counts are dependent; bins chosen after looking at the data. Chi-square approximation to null distribution of χ^2_{MC} can be poor, but can estimate P-values by simulation.

Kolmogorov-Smirnov Test (KS)

- Test whether, conditional on the number of events, re-scaled times are iid $U[0,1]$.

$$\text{KS statistic } (U[0,1] \text{ null}): \; D_n = \sup_t \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(t_i \leq t) - t \right| .$$

- Doesn't require estimating parameters or ad hoc N_w, K^-, K^+, $\hat{\lambda}$.

Kolmogorov-Smirnov Test (KS)

- Test whether, conditional on the number of events, re-scaled times are iid $U[0,1]$.

$$\text{KS statistic } (U[0,1] \text{ null}): \; D_n = \sup_t \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(t_i \leq t) - t \right| .$$

- Doesn't require estimating parameters or ad hoc N_w, K^-, K^+, $\hat{\lambda}$.

Power, limitations

- KS: long-term rate variations
- PD & BZ: heterogeneity across intervals
- MC: departure from Poisson distribution across intervals
- PD, BZ & MC insensitive to the order of the intervals: rearrangements don't matter
  For instance, the statistics have same value for
  $(N_k) = (3, 1, 0, 2, 0, 4, 1, 0)$ as for $(N_k) = (0, 0, 0, 1, 1, 2, 3, 4)$.
  But latter hardly looks Poisson.
- PD, BZ, & KS are insensitive to under-dispersion, e.g., equispaced events. (can rectify using two-sided versions)
- None uses spatial information!
- Moreover, many *ad hoc* choices for PD, BZ, & especially MC
- MC internally inconsistent; chi-square approximation can be especially poor.

Tests on simulated data

Process	KS power	MC power
Heterogeneous Poisson	1	0.1658
Gamma renewal	0.0009	1

Estimated power of level-0.05 tests of homogeneous Poisson null hypothesis from 10,000 simulations. “Heterogeneous Poisson”: rate 0.25 per ten days for 20 years, then rate 0.5 per ten days for 20 years. “Gamma renewal”: inter-event times iid gamma with shape 2 and rate 1. MC test uses 10-day intervals, 4 categories, $d = 2$ degrees of freedom.

Declustering Methods

- Window-based methods
  - Main-shock window: punch hole in catalog near each “main shock”
  - Linked window: every event has a window.
    Clusters are maximal sets of events such that each is in the window of some other event in the group.
    Replace cluster by single event: first, largest, “equivalent”
- Stochastic methods: use chance to decide which events to keep
- Other methods (e.g., waveform similarity)
- Straw man: deTest.

Declustering Methods

- Window-based methods
  - Main-shock window: punch hole in catalog near each "main shock"
  - Linked window: every event has a window.
    Clusters are maximal sets of events such that each is in the window of some other event in the group.
    Replace cluster by single event: first, largest, "equivalent"
- Stochastic methods: use chance to decide which events to keep
- Other methods (e.g., waveform similarity)
- Straw man: deTest.

Declustering Methods

- Window-based methods
  - Main-shock window: punch hole in catalog near each “main shock”
  - Linked window: every event has a window.
    Clusters are maximal sets of events such that each is in the window of some other event in the group.
    Replace cluster by single event: first, largest, “equivalent”
- Stochastic methods: use chance to decide which events to keep
- Other methods (e.g., waveform similarity)
- Straw man: deTest.

Declustering Methods

- Window-based methods
  - Main-shock window: punch hole in catalog near each “main shock”
  - Linked window: every event has a window.
    Clusters are maximal sets of events such that each is in the window of some other event in the group.
    Replace cluster by single event: first, largest, “equivalent”
- Stochastic methods: use chance to decide which events to keep
- Other methods (e.g., waveform similarity)
- Straw man: deTest.

deTest Declustering

No physical basis. Not intended to be a method for declustering catalogs for seismological purposes, only a “straw man” to show:

1. A declustered catalog can have rather more events than window-based declustering methods leave, and still pass tests for temporally homogeneous Poisson behavior.
2. Using spatial and temporal data by testing for conditional exchangeability of times given the locations (described below) can be more powerful than testing only for temporal homogeneity.

deTest details

K is given. Declustering a catalog to make the result pass the MC, CC, or BZ test is constrained by the number of intervals among the K in the original catalog that have no events: declustering can delete events but not add them.

Number of intervals with 0 events gives an implicit estimate of the rate of a Poisson process that the declustered catalog can be coerced to fit well:
If seismicity followed a homogenous Poisson process with theoretical rate λ events per interval, the chance that an interval would contain no events is $e^{-\lambda}$. So,

$$\hat{\lambda} \equiv -\log\left(\frac{\text{\# intervals with no events}}{K}\right)$$

is a natural estimate of the rate of events per interval.

deTest Strategy

deTest tries to construct a catalog with about this rate that passes all four tests described above (MC, CC, BZ, and KS).

If seismicity followed a homogeneous Poisson distribution with theoretical rate per interval $\hat{\lambda}$, then the expected number of intervals in the declustered catalog with at least c events would be

$$G_c \equiv \sum_{i=c}^{\infty} K \times \frac{\hat{\lambda}^i e^{-\hat{\lambda}}}{i!}.$$

deTest constructs a catalog in which $[G_c]$ intervals contain c or more events, where $[x]$ denotes the integer closest to x.

deTest Algorithm

Start with empty catalog. Add events until the result has approximately the correct expected number of intervals with each number of events, then remove events until pass the KS test.

1. Count the events in the raw catalog in each of the N_w intervals.
2. Define $\hat{\lambda}$ and G_c as above.
3. Let $c = 1$. From each interval in the raw catalog that has at least one event, include one event selected at random from that interval.
4. Let $c \leftarrow c + 1$. If $[G_c] = 0$, go to step 6. Otherwise, go to step 5.

5. Add events until $[G_c]$ intervals have at least c events, while keeping the KS statistic small: Let N_t be the number of events in the construct catalog that have occurred by time t. Find element t_m of the set $t \in \{T/K, 2T/K, \ldots, T\}$ at which $N_t/N_T - t/T$ is minimized. Find the intervals that
   a. contain $c - 1$ events in the current declustered catalog;
   b. contain at least c events in the raw catalog.

   From this set, select the interval prior to t_m but closest to t_m (If no interval is prior to t_m, choose the first interval in the set.) Choose event randomly from events in the selected interval that have not yet in the declustered catalog; add that event to the declustered catalog. Repeat until $[G_c]$ intervals contain c events, then return to step (4).

6. Find the KS P-value p_{KS}. If $p_{KS} < \alpha$, find a time t at which the empirical cdf differs maximally from the uniform cdf. Either t is infinitesimally before an event or t is the time of one or more events. If t is just before an event, t/T is larger than the empirical cdf: Delete the event after time t at which the empirical cdf minus the uniform cdf is largest. If t is the time of an event, the empirical cdf at t is larger than t/T: Delete an event at time t. Repeat until $P_{KS} > \alpha$.

Are "main events" Poisson in time?

Gardner & Knopoff, 1974:
"Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?"

Abstract: "Yes."

Statistical test: MC

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test. Shrug.

MC test on a number of declustered catalogs, including a catalog of 1,751 $M \geq 3.8$ events in Southern California, 1932–1971.

Similar to 1932–1971 SCEC catalog, not identical (1,556 $M \geq 3.8$ events)

Declustered: 503 events. 10-day intervals. $d = 2$ degrees of freedom. Don't give B; don't explain how λ estimated.

Are “main events” Poisson in time?

Gardner & Knopoff, 1974:
“Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?”

Abstract: “Yes.”

Statistical test: MC

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test. Shrug.

MC test on a number of declustered catalogs, including a catalog of 1,751 $M \geq 3.8$ events in Southern California, 1932–1971.

Similar to 1932–1971 SCEC catalog, not identical (1,556 $M \geq 3.8$ events)

Declustered: 503 events. 10-day intervals. $d = 2$ degrees of freedom. Don't give B; don't explain how λ estimated.

Are “main events” Poisson in time?

Gardner & Knopoff, 1974:
“Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?”

Abstract: “Yes.”

Statistical test: MC

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test. Shrug.

MC test on a number of declustered catalogs, including a catalog of 1,751 $M \geq 3.8$ events in Southern California, 1932–1971.

Similar to 1932–1971 SCEC catalog, not identical (1,556 $M \geq 3.8$ events)

Declustered: 503 events. 10-day intervals. $d = 2$ degrees of freedom. Don’t give B; don’t explain how λ estimated.

Are “main events” Poisson in time?

Gardner & Knopoff, 1974:
“Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?”

Abstract: “Yes.”

Statistical test: MC

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test. Shrug.

MC test on a number of declustered catalogs, including a catalog of 1,751 $M \geq 3.8$ events in Southern California, 1932–1971.

Similar to 1932–1971 SCEC catalog, not identical (1,556 $M \geq 3.8$ events)

Declustered: 503 events. 10-day intervals. $d = 2$ degrees of freedom. Don’t give B; don’t explain how λ estimated.

Are “main events” Poisson in time?

Gardner & Knopoff, 1974:
“Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?”

Abstract: “Yes.”

Statistical test: MC

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test.
Shrug.

MC test on a number of declustered catalogs, including a catalog of 1,751 $M \geq 3.8$ events in Southern California, 1932–1971.

Similar to 1932–1971 SCEC catalog, not identical (1,556 $M \geq 3.8$ events)

Declustered: 503 events. 10-day intervals. $d = 2$ degrees of freedom. Don’t give B; don’t explain how λ estimated.

Tests on SCEC data

- GKl: Remove every event in the window of some other event.
- GKlb: Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- Method GKm: Consider the events in chronological order. If the *i*th event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the *i*th event, delete the *i*th event. Otherwise, retain the *i*th event.
- Rl: Reasenberg's (1985) linked-window method
- dT: deTest, described above. Keep events from catalog deliberately to pass MC and KS tests.

Tests on SCEC data

- GKl: Remove every event in the window of some other event.
- GKlb: Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- Method GKm: Consider the events in chronological order. If the ith event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the ith event, delete the ith event. Otherwise, retain the ith event.
- Rl: Reasenberg's (1985) linked-window method
- dT: deTest, described above. Keep events from catalog deliberately to pass MC and KS tests.

Tests on SCEC data

- GKl: Remove every event in the window of some other event.
- GKlb: Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- Method GKm: Consider the events in chronological order. If the *i*th event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the *i*th event, delete the *i*th event. Otherwise, retain the *i*th event.
- Rl: Reasenberg's (1985) linked-window method
- dT: deTest, described above. Keep events from catalog deliberately to pass MC and KS tests.

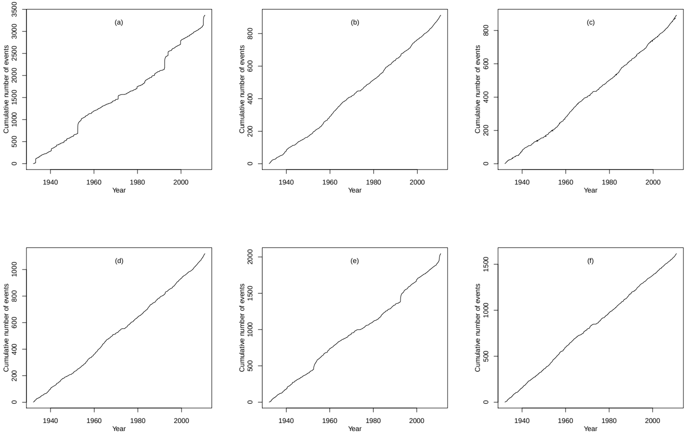
Tests on SCEC data

- GKl: Remove every event in the window of some other event.
- GKlb: Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- Method GKm: Consider the events in chronological order. If the *i*th event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the *i*th event, delete the *i*th event. Otherwise, retain the *i*th event.
- Rl: Reasenberg's (1985) linked-window method
- dT: deTest, described above. Keep events from catalog deliberately to pass MC and KS tests.

Tests on SCEC data

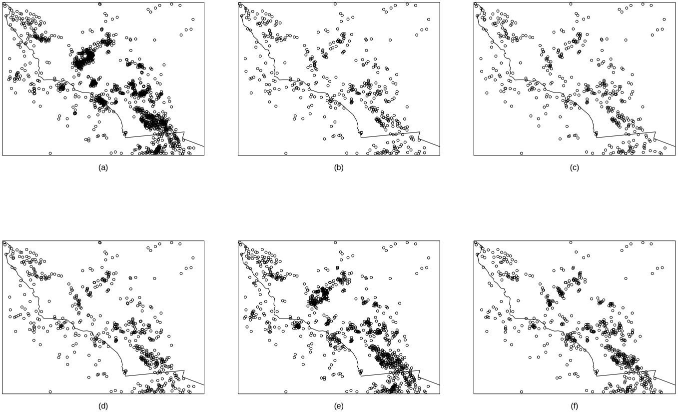
- GKl: Remove every event in the window of some other event.
- GKlb: Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- Method GKm: Consider the events in chronological order. If the *i*th event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the *i*th event, delete the *i*th event. Otherwise, retain the *i*th event.
- Rl: Reasenberg's (1985) linked-window method
- dT: deTest, described above. Keep events from catalog deliberately to pass MC and KS tests.

Cumulative temporal distribution of events, 1932–2010



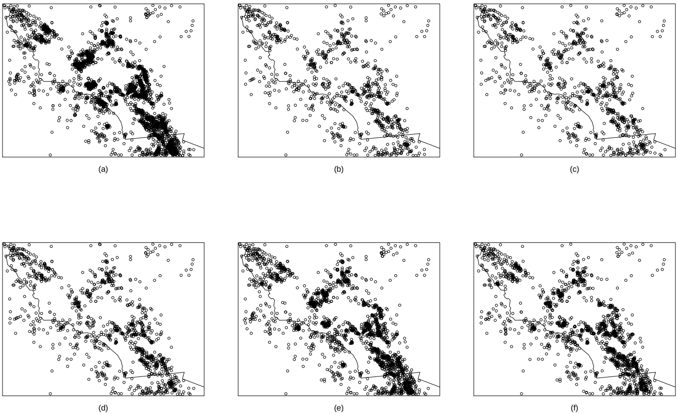
(a) All 3,368 $M \geq 3.8$ events. (b) 913 GKl events. (c) 892 GKLb events. (d) 1,120 GKm events. (e) 2,046 Rl events. (f) 1,615 dT events.

SCEC $M \geq 3.8$, 1932–1971



(a) 1,556 original; (b) 437 GKl; (c) 424 GKlb. (d) 544 GKm.
(e) 985 Rl. (f) 608 dT.

SCEC $M \geq 3.8$, 1932–2010



(a): 3,368 original; (b) 913 GKl; (c) 892 GKlb; (d) 1,120 GKm; (e) 2,046 Rl; (f) 1,615 dT.

Exchangeability of times

- For SITHP, marginal distribution of times is Poisson, so when temporal test rejects, implicitly rejects SITHP.
- For SITHPs, two events can be arbitrarily close. Window declustering imposes minimum spacing, so can't be SITHP.
- For SITHPs, conditional on the number of events, event locations in space-time are iid with probability density proportional to the space-time rate. Conditional on the locations, the marginal distribution of times is iid, hence exchangeable.

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \stackrel{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\}$$

for all permutations $\pi \in \Pi$.
($\stackrel{d}{=}$ means “has the same probability distribution as”)

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \stackrel{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\}$$

for all permutations $\pi \in \Pi$.
($\stackrel{d}{=}$ means "has the same probability distribution as")

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \stackrel{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\}$$

for all permutations $\pi \in \Pi$.
($\stackrel{d}{=}$ means “has the same probability distribution as”)

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \overset{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\}$$

for all permutations $\pi \in \Pi$.
($\overset{d}{=}$ means "has the same probability distribution as")

Exchangeability, contd.

- SITHP has exchangeable times.
- If events close in space tend to be close in time—the kind of clustering real seismicity exhibits—times not exchangeable.
- If events close in space tend to be distant in time—e.g., from window methods for declustering—times not exchangeable.

Exchangeability, contd.

- SITHP has exchangeable times.
- If events close in space tend to be close in time—the kind of clustering real seismicity exhibits—times not exchangeable.
- If events close in space tend to be distant in time—e.g., from window methods for declustering—times not exchangeable.

Exchangeability, contd.

- SITHP has exchangeable times.
- If events close in space tend to be close in time—the kind of clustering real seismicity exhibits—times not exchangeable.
- If events close in space tend to be distant in time—e.g., from window methods for declustering—times not exchangeable.

Permutation test set up

- Null of SITHP satisfies group invariance: If null holds, equally likely to observe same locations with any permutation of the times.
- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V = \{(x, y, t) \in R^3 : x \leq x_0 \text{ and } y \leq y_0 \text{ and } t \leq t_0\}.$$

- **V**: set of all lower-left quadrants.

Permutation test set up

- Null of SITHP satisfies group invariance: If null holds, equally likely to observe same locations with any permutation of the times.
- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V = \{(x, y, t) \in R^3 : x \le x_0 \text{ and } y \le y_0 \text{ and } t \le t_0\}.$$

- **V**: set of all lower-left quadrants.

Permutation test set up

- Null of SITHP satisfies group invariance: If null holds, equally likely to observe same locations with any permutation of the times.
- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V = \{(x, y, t) \in R^3 : x \leq x_0 \text{ and } y \leq y_0 \text{ and } t \leq t_0\}.$$

- **V**: set of all lower-left quadrants.

Permutation test set up

- Null of SITHP satisfies group invariance: If null holds, equally likely to observe same locations with any permutation of the times.
- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V = \{(x, y, t) \in R^3 : x \leq x_0 \text{ and } y \leq y_0 \text{ and } t \leq t_0\}.$$

- **V**: set of all lower-left quadrants.

Test statistic

$$\phi(\{(x_j, y_j, t_j)\}_{j=1}^{n}) \equiv \sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$$

- Supremum of the difference between probabilities of sets generalizes KS statistic to three dimensions.
- Suffices to search a finite subset of $\mathbf{V}$.
  Can sample at random from that finite subset for efficiency.
- Estimate P-value by simulating from $\tau(\hat{P}_n)$—permuting the times (Romano)

Test statistic

$$\phi(\{(x_j, y_j, t_j)\}_{j=1}^{n}) \equiv \sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$$

- Supremum of the difference between probabilities of sets generalizes KS statistic to three dimensions.
- Suffices to search a finite subset of **V**.
  Can sample at random from that finite subset for efficiency.
- Estimate P-value by simulating from $\tau(\hat{P}_n)$—permuting the times (Romano)

Test statistic

$$\phi(\{(x_j, y_j, t_j)\}_{j=1}^{n}) \equiv \sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$$

- Supremum of the difference between probabilities of sets generalizes KS statistic to three dimensions.
- Suffices to search a finite subset of **V**.
  Can sample at random from that finite subset for efficiency.
- Estimate *P*-value by simulating from $\tau(\hat{P}_n)$—permuting the times (Romano)

Years	Mag (events)	Meth	n	MC χ^2	MC Sim	PD	BZ	KS	Romano P	Reject? Time	Reject? Space-time
32–71	3.8 (1,556)	GKl	437	0.087	0.089	0.069	0.096	0.011	0.005	Yes	Yes
		GKlb	424	0.636	0.656	0.064	0.108	0.006	0.000	Yes	Yes
		GKm	544	0	0	0	0	0.021	0.069	Yes	No
		Rl	985	0	0	0	0	0.003	0	Yes	Yes
		dT	608	0.351	0.353	0.482	0.618	0.054	0.001	No	Yes
	4.0 (1,047)	GKl	296	0.809	0.824	0.304	0.344	0.562	0.348	No	No
		GKlb	286	0.903	0.927	0.364	0.385	0.470	0.452	No	No
		GKm	369	<0.001	<0.001	0	0	0.540	0.504	Yes	No
		Rl	659	0	0	0	0	0.001	0	Yes	Yes
		dT	417	0.138	0.134	0.248	0.402	0.051	0	No	Yes
32–10	3.8 (3,368)	GKl	913	0.815	0.817	0.080	0.197	0.011	0.214	Yes	No
		GKlb	892	0.855	0.855	0.141	0.204	0.005	0.256	Yes	No
		GKm	1120	0	0	0	0	0.032	0.006	Yes	Yes
		Rl	2046	0	0	0	0	0	0	Yes	Yes
		dT	1615	0.999	1.000	0.463	0.466	0.439	0	No	Yes
	4.0 (2,169)	GKl	606	0.419	0.421	0.347	0.529	0.138	0.247	No	No
		GKlb	592	0.758	0.768	0.442	0.500	0.137	0.251	No	No
		GKm	739	0	0	0	0	0.252	0.023	Yes	Yes
		Rl	1333	0	0	0	0	0	0	Yes	Yes
		dT	1049	0.995	0.999	0.463	0.465	0.340	0	No	Yes

Discussion: Seismology

- Regional declustered catalogs generally don't look Poisson in time.
- Window-declustered catalogs *can't* be Poisson in space-time.
- Window-declustered catalogs generally don't seem to have exchangeable times, necessary condition for Poisson.
- No physics definition of foreshock, main shock, aftershock.

Discussion: Statistics

- The test matters. What's the scientific question?
- Power of tests varies dramatically across alternatives.
- Easy to make declustering method pass tests if you try.
- Novel test for exchangeability of times given locations and times.