On survival analyses of the timing of cognitive processes

Pablo Gomez, Javier Breithaupt, Manuel Perea & Jeff Rouder

DePaul U, Universitat de Valencia, U of Missouri

Latencies

- Understanding the micro-structure of laboratory tasks informs us about the components of cognitive function.
- Several tasks utilize latencies as the main dependent variable.
- RTs, fixation times, naming, are examples of latency based measurements.

What we do with latency measurements?

- Often, effects are explored though null-hypothesis testing on the mean latency.
- This approach might be appropriate if the goal is to obtain evidence for the existence of a phenomenon *per se*.
- However, it might not be as informative if we are trying to understand specifically what component of processing is affected by a manipulation.

What to do...

There are methods to examine latency data, and each one of them requires you to make some level of theoretical assumption.

1. *Mean RT*. No or little theoretical commitment, but misses the distributional information; however, with clever experiments (e.g., masked priming) researcher can make assumptions about the timing and locus of effects.

What to do...

There are methods to examine latency data, and each one of them requires you to make some level of theoretical assumption.

1. *Mean RT*. No or little theoretical commitment, but misses the distributional information; however, with clever experiments (e.g., masked priming) researcher can make assumptions about the timing and locus of effects.
2. *Functional form analyses.* It assumes that latencies can be described by distributions with known functional forms: ex-Gaussian, Weibull, etc.

What to do...

There are methods to examine latency data, and each one of them requires you to make some level of theoretical assumption.

1. *Mean RT*. No or little theoretical commitment, but misses the distributional information; however, with clever experiments (e.g., masked priming) researcher can make assumptions about the timing and locus of effects.
2. *Functional form analyses.* It assumes that latencies can be described by distributions with known functional forms: ex-Gaussian, Weibull, etc.
3. *Process Models.* A process model provides with a mechanism that generates the latency data. One fits the model to the data, and then obtains parameters.

Survival analyses

Here we explore an alternative to the three methods just mentioned: survival analyses.

The promise of survival analyses

1. It goes beyond the mean RT, as it explores the full distributions of latencies.
2. It does not make assumptions about functional form, and is not subject to misspecification.
3. It does not make a theoretical commitment to a process model.

The promise of survival analyses

In short, it might be right on the *sweet spot* of data analysis.

What is survival?

- $S(t) = Pr(T > t)$
- For a given time t, the proportion of responses with $latency > t$ are the proportion survival time at time t.
- At $t = 0$, $S(0) = 1$, and at longest latency, $S(t) \to 0$ as $t \to \infty$

In the past...

- There have been attempts to use survival analyses and hazard functions; see Van Zandt (2002)
- "Serious hazard function analysis would use samples of at least a few hundred observations"
- Estimating the tail is specially difficult.

Let's consider a manipulation with two conditions

- A research question: Can we learn anything about the temporal dynamics of different components of processing?

Let's consider a manipulation with two conditions

- A research question: Can we learn anything about the temporal dynamics of different components of processing?
- Divergence point corresponds to the shortest latency value at which the a manipulation has a significant impact.

Let's consider a manipulation with two conditions

- A research question: Can we learn anything about the temporal dynamics of different components of processing?
- Divergence point corresponds to the shortest latency value at which the a manipulation has a significant impact.
- Sheridan (2013); Reingold et al. (2012) for first fixation durations.
- 1-ms time bins: the survival curve was computed separately for each condition and for each participant, and then averaged across participants.

Divergence point estimation. Boot strapping method

- On each iteration, the latencies for each participant in each condition was randomly re-sampled with replacement.

Divergence point estimation. Boot strapping method

- On each iteration, the latencies for each participant in each condition was randomly re-sampled with replacement.
- Each participants' survival curves were then computed and averaged (across subjects).

Divergence point estimation. Boot strapping method

- On each iteration, the latencies for each participant in each condition was randomly re-sampled with replacement.
- Each participants' survival curves were then computed and averaged (across subjects).
- Next, for each time bin t the $\Delta_{t,i}$ are computed (i^{th}: iteration from 1 to 10,000), and then sorted.

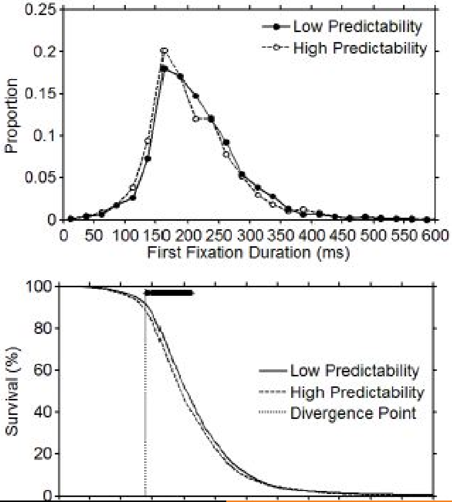
Divergence point estimation. Boot strapping method

- On each iteration, the latencies for each participant in each condition was randomly re-sampled with replacement.
- Each participants' survival curves were then computed and averaged (across subjects).
- Next, for each time bin t the $\Delta_{t,i}$ are computed (i^{th}: iteration from 1 to 10,000), and then sorted.
- The range between the 5th and the 9,995th value becomes $CI_{\Delta(t)}$

Divergence point estimation. Boot strapping method

- On each iteration, the latencies for each participant in each condition was randomly re-sampled with replacement.
- Each participants' survival curves were then computed and averaged (across subjects).
- Next, for each time bin t the $\Delta_{t,i}$ are computed (i^{th}: iteration from 1 to 10,000), and then sorted.
- The range between the 5th and the 9,995th value becomes $CI_{\Delta(t)}$
- Divergence point between conditions: the t at which $CI_{\Delta(t)} doesnotinclude 0$

Sheridan, 2013



Divergence point

- Clearly, if this method works there could be many applications for it.
- Sheridan and cols have presented analyses of eye movement data using this method.

Let's cut to the chase

Estimating divergence points in latency distributions is conceptually flawed.

Assumption

The two conditions share a process, and this is followed by a subsequent processing component in which the two conditions differ (diverge) on the timing and/or the processing cost.

Realistic assumption?

Although one could think of somewhat contrived situations in which this type of serial, staged process might occur (see Carreiras, Armstrong, Perea, & Frost, 2014, for a recent review against this type of formulation), the latencies, even at their shortest duration, would be a reflection of the **ending times** of the second process.

In short

Latency distributions reflect aggregation across processes, across trials, and even across participants.

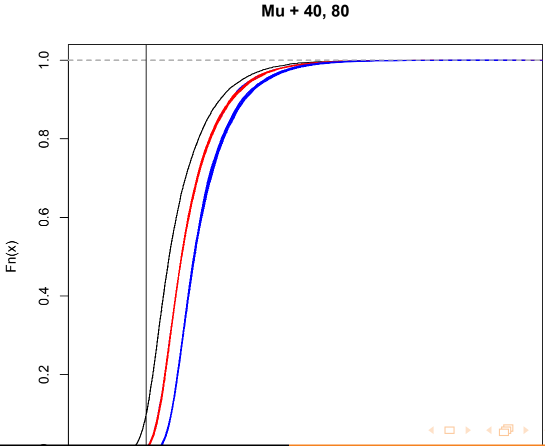
How do we know this?

Latency measurements show stochastic dominance.

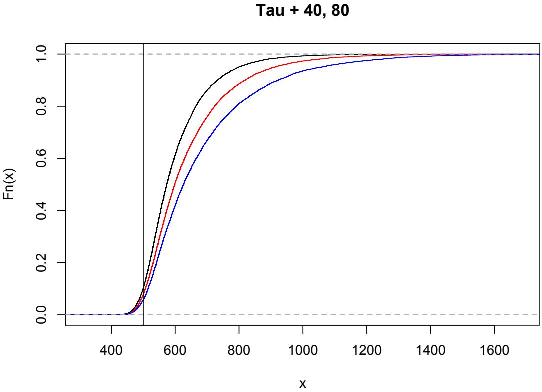
Stochastic dominance

Stochastic dominance refers to the probability of observations smaller than x being greater for one variable than the other for all values of x (see Heathcote, Brown, Wagenmakers, & Eidels, 2010).

Example,cumulative density functions generated with an ex-Gaussian distribution in which there are effects on μ



Example,cumulative density functions generated with an ex-Gaussian distribution in which there are effects on τ



Rationale

What are the consequences of applying this method to data in which there is stochastic dominance?

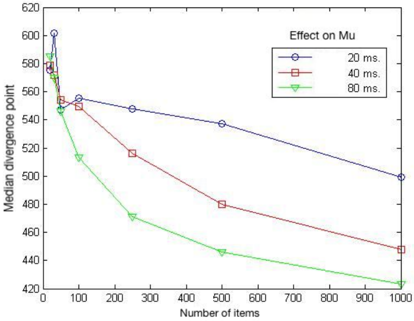
- Generate data from an ex-gaussian distribution assuming that the experimental effect is either in the μ or τ parameters of the ex-gaussian.
- Create latencies manipulating the number of hypothetical trials per condition.
- We apply the bootstrapping method.
- We explore if the method recovers the properties of the data generating strategy.

Effects on μ

- Data generated from an ex-gaussian with $\mu = 541$; $\sigma = 68$; $\tau = 115$; $\Delta_\mu = 20, 40, 80$.
- Shift in distribution.... in other words, stochastic dominance, divergence happens at shortest latencies!

Effects on μ

A biased estimation of divergence point

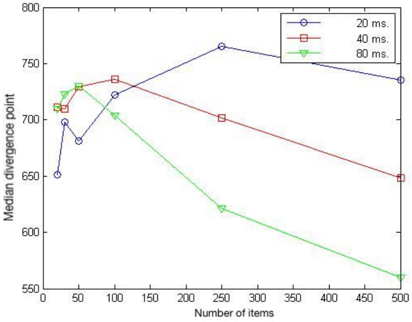


Effects on τ with large τ values

- Data generated from an ex-gaussian with $\mu = 541$; $\sigma = 68$; $\tau = 115$; $\Delta_\tau = 20, 40, 80$.
- Change in tail... in other words, stochastic dominance, divergence happens at shortest latencies!

Effects on τ with large τ values

A biased estimation of divergence point

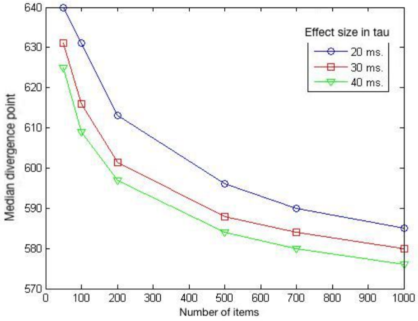


Effects on τ with small τ values

- Data generated from an ex-gaussian with $\mu = 541$; $\sigma = 68$; $\tau = 20$; $\Delta_\tau = 20, 40, 80$.
- Change in tail : The method should give us a divergence point later on.

Effects on τ with small τ values

A biased estimation of divergence point



Summary of results

- With number of items within the standard cognitive psychology experiment, the method severely overestimates the point in time of the divergence.
- As it stands, the method provides with an output that is mostly related to the number of items per condition.

Take home message (I)

Latency measurements tend to exhibit stochastic dominance between experimental conditions, and hence the divergence point would be at the leading edge of the latency distribution regardless of other distributional differences.

Take home message (II)

Furthermore, if the method is applied to data, an estimate of the divergence point will be provided by the method. This estimate will be affected heavily by the number of observations. In short, our exploration of the method forces us to conclude that it is not advisable to utilize it when analyzing latency data.

Thank you!