Estimation of Travel Time Distributions and Travel Time Derivatives

Ke Wan

Operations Research Financial Engineering, Princeton University

June 8, 2014
Outline

1 Introduction to Transportation Risk Management

2 Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)

3 Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
   - Modeling Travel Time Changes as Continuous Time ARMA \((p, q)\) Processes
   - Pricing Travel Time Derivatives

4 Conclusion and Future Work
Complex systems with substantially varied performance

- The cumulative effects of numerous individual behaviors;
- A fundamental performance measurement of the transportation system is travel time.
- Uncertainties in travel time have a substantial impact on travelers and related business.

Goal: Develop analytical methods for the estimation and hedging of travel time related risk

- To characterize the uncertainties of travel time and generate routing decisions, we need:
  - To estimate travel time distributions;
  - To measure risk and routing given the distribution.
- To design innovative value pricing schemes and hedge economic risk via financial derivatives.
Definition and Measurement of Travel Time

- Travel time is the period of time spent by an entity traveling between two points in a road network, which is measured and indexed in specific ways.

Figure: A sample road network and a typical link travel time distribution (numbers are in seconds)

- Travel time may change due to differences in:
  - Site-based v.s. Vehicle based based measurements;
  - Individual v.s. Average measurements;
  - Index method: Arrival time v.s. Departure time;
  - Base network: physical v.s. monument-to-monument.
Major challenge

- Estimate travel time distribution for Origin-Destination pairs based on segment-wise travel time.
  - There are $O(N^2)$ O-D pairs given $N$ points of interest.

- Estimate O-D path travel time based on $O(N)$ segments through convolution.
Outline

1 Introduction to Transportation Risk Management

2 Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)

3 Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
   - Modeling Travel Time Changes as Continuous Time ARMA($p, q$) Processes
   - Pricing Travel Time Derivatives

4 Conclusion and Future Work
Measurement and Model for Travel Time

- How do we estimate a travel time distribution?
- Review of existing research:
  - Link travel time estimation:
    - Kalman filter (Chu (2005) [5])
    - Artificial neural network models (Mark (2004) [13]),
    - Gaussian mixtures applied to a single link (Park (2011) [15]),
  - Path travel time estimation:
    - Assumption of independence or conditional independence:
      (Timothy (2009) [12], Herring (2010) [11])
    - Joint Gaussian distributions: (Dailey (2000) [7], Zhan (2013) [21]).
- Measurements for travel time:
  - Based on random travel time measurements from Copilot GPS devices;
    - Heavy-tailed
    - Unevenly distributed & sparse in time and space
  - Based on individual trip travel time on the given path;
  - Indexed by departure time for prediction purposes.
The Definition of Copulas

Let \( P \) be a conditional bivariate distribution function with continuous margins \( F_X \) and \( F_Y \), and let \( \mathcal{F} \) be some conditioning set. There then exists a unique conditional copula \( C : [0, 1] \times [0, 1] \) such that (Sklar (1959) [16]):

\[
P(x, y|\mathcal{F}) = C(F_X(x|\mathcal{F}), F_Y(y|\mathcal{F})|\mathcal{F}), \forall x, y \in \mathbb{R}
\]  

(1)

Advantages of the copula method:
- Heavy-tailed joint distributions can be modeled;
- Marginal travel time distributions and their dependent structure can be studied separately;
- Copulas can be calibrated to data sets that are sparse and unevenly distributed.

Upper tail dependence of different copulas (Nelson (2006) [14]): How likely is there a congestion in a nearby link if one observes it at current place?

<table>
<thead>
<tr>
<th>Name</th>
<th>Lower Tail</th>
<th>Upper Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian Copula</td>
<td>( \lambda_L = 0 )  ( \lambda_U = 0 )</td>
<td></td>
</tr>
<tr>
<td>T Copula</td>
<td>( \lambda_L = 2t_{\mu+1}(-\sqrt{\mu+1}\sqrt{\frac{1-\rho}{1+\rho}}) ) ( \lambda_U = \lambda_L )</td>
<td></td>
</tr>
<tr>
<td>BB1</td>
<td>( \lambda_L = 2^{-\frac{1}{\sigma}} ) ( \lambda_U = 2 - 2^{\frac{1}{\sigma}} )</td>
<td></td>
</tr>
</tbody>
</table>
Two-step Estimation and Goodness-of-fit for Copulas

- **Step 1:** Estimate marginal distributions by kernel methods;
- **Step 2:** Copula estimation (White (1994) [18]):

<table>
<thead>
<tr>
<th>link</th>
<th>Normal</th>
<th>T</th>
<th>BB1</th>
<th>Clayton</th>
<th>Gumbel</th>
<th>Frank</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>-80.1444</td>
<td>-84.5242</td>
<td>-85.9260</td>
<td>-71.5360</td>
<td>-76.0585</td>
<td>-72.3326</td>
</tr>
<tr>
<td>p-values</td>
<td>N/A</td>
<td>N/A</td>
<td>0.8110</td>
<td>0.1308</td>
<td>0.0058</td>
<td>0.0350</td>
</tr>
</tbody>
</table>

- Aggregate the travel times of two nearby links along a path using BB1 copula.

**Figure:** Left: Network topology; Center: Estimated BB1 copula; Right: travel time distribution of path AB-BE Estimated: Red; Empirical: Black
Outline

1. Introduction to Transportation Risk Management

2. Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)

3. Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
   - Modeling Travel Time Changes as Continuous Time ARMA($p, q$) Processes
   - Pricing Travel Time Derivatives

4. Conclusion and Future Work
Estimation of the Path Travel Time Distribution

- Estimate the path travel time distribution using data from individual links.

- Issues to consider:
  - Stationarity: path scenarios;
  - Dependence: a mixture of copula based models;
  - Heavy tail: non-parametric estimation for marginal distributions;
Path travel time aggregation: Path Scenario

Definition

A path scenario describes the overall traffic conditions of a given path. Given a path scenario, the travel time for sections along the path is subject to a stationary joint distribution.

Observable: Through experience, modelers can categorize observed travel time data into predefined path scenarios.

Unobservable: When path scenarios are termed unobservable, modelers extract path scenarios from the data.
A GCMM is a mixture of joint distribution functions, each of which contains a Gaussian copula

\[ F(X|\pi) = \sum_{k=1}^{K} \pi_k \int_{-\infty}^{Y_{k1}} \cdots \int_{-\infty}^{Y_{kd}} \frac{1}{(2\pi)^{n/2}P_k^{1/2}} \exp\left(-\frac{1}{2}Y_k^TP_kY_k\right) dY \]  

where

- \( x = [x_1 \ldots x_d] \) is the marginal observation.
- \( Y_k = [Y_{1d} \ldots Y_{kd}] \) is the vector of the transferred data.
- \( Y_{kd} = \Psi^{-1}(F_{kd}(x_d)) \) is the d-th dimension of the transferred data.
- \( Z_{kd} = f_{kd}(x_d) = \frac{\partial F_{kd}}{\partial x}(x_d) \) is the density of the marginal distribution.
- \( \pi_k \) is the weight to the \( k \)-th copula.

Its density is given by

\[ f(X|\pi) = \sum_{k=1}^{K} \pi_k \frac{1}{(2\pi)^{n/2}P_k^{1/2}} \exp\left(-\frac{1}{2}Y_k^TP_kY_k\right) \prod_{d=1}^{D} \frac{Z_{kd}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(Y_{kd})^2\right) \]
Identification of Path Scenarios

- Regime identification with the help of cluster analysis:

![Hierarchy cluster for travel time data of a three-segment path](image)

*Figure: Hierarchy cluster for travel time data of a three-segment path (Left: Original data; Center: top 30 clusters; Right: Number of observations across clusters)*

- Regimes identification and factor analysis
  - Congestion level (Ranges of travel time)
  - Time of day
  - Weather (Normal / Rainy)
  - Event (Yes / No)
The estimation procedure includes:

- Data categorization;
- Estimation of marginal distribution in each scenario;
- Estimation of copula parameters via Lasso method on the transformed data.

\[
\max_X \log(\det X) - tr(\Sigma^T X) - v|X|_1
\]  

(4)

where $|X|_1$ is the sum of the terms in $X$ (Friedman (2008) [9]) and $\Sigma$ is the covariance matrix for $\Psi^{-1}(F(X))$

- Robust estimation;
- Shrink between noise and bias;

Advantage: A straightforward estimation procedure.
An estimation using path scenarios which is identified by cluster analysis:

Figure: Estimated path travel time distribution: Based on Major Clusters: Green; With Outliers: Red; Empirical: Black
Limitation to observable path scenarios:
- Identification of scenarios requires experience;
- Treatment of outliers can be improved;
- Factor sensitivity may not be high based on regression analysis;

GCMM based on Unobservable Path Scenarios:
- Expectation Maximum algorithm (Dempster (1977) [8]);
- Gaussian mixture models (GMM) (Xu (1996) [20]).

Intuition: Can we introduce copulas into Gaussian mixture models?
- A mixture of joint distributions, each with a Gaussian copula;
- Copulas and weights are estimated using the EM algorithm;
- Each marginal distribution is estimated through nonparametric methods.
GCMM updates mixtures of copulas and the marginal distributions separately:

**Figure:** Comparison of GMM and GCMM base model: n: data index; m: iteration index; k: copula index; i: dimension
GCMM based on Unobservable Path Scenarios

Advantages:

- Achieve more flexible fitting by updating copula structures and marginal distributions separately;
- Reduce number of clusters compared to Gaussian mixture models (GMM) (Xu (1996) [20]);
- One version may further utilize unsynchronized travel time observations on each link.

Figure: Comparison of GMM and GCMM
Expectation Maximum Algorithm for GCMM 2

- **Expectation Step:**

\[
D_{nk}^m = \prod_{i=1}^{D} \frac{Z_{n,ki}^m}{\sqrt{2\pi}} \exp\left(\frac{1}{2} (Y_{n,ki}^m)^2 \right) \tag{5}
\]

\[
r_{nk}^m = \frac{\pi_k \frac{D_{n,k}^m}{|P_k^m|^{1/2}} \exp\left( -\frac{1}{2} (Y_{nk}^m)^T P_k^m, -1 Y_{nk}^m \right)}{\sum_{j=1}^{K} \pi_j \frac{D_{n,j}^m}{|P_j^m|^{1/2}} \exp\left( -\frac{1}{2} (Y_{nj}^m)^T P_j^m, -1 Y_{nj}^m \right)} \tag{6}
\]

- **Maximum Step:**

\[
\pi_k^m = \frac{\sum_{n=1}^{N} r_{nk}^m}{N} \tag{7}
\]

\[
P_k^m = \frac{\sum_{n=1}^{N} r_{nk}^m Y_{nk}^m (Y_{nk}^m)^T}{\sum_{n=1}^{N} r_{nk}^m} \tag{8}
\]

\[
F_{ki}^m(y) = \frac{\sum_{n} r_{nk}^m 1_{x_{ni} \leq y}}{\sum_{n} r_{nk}^m} \tag{9}
\]

\forall k\text{-th copula, } i\text{-th dimension}
Theoretical properties

- Bounded likelihood function;

\[
L = \sum_{n=1}^{N} \ln\left(\sum_{k=1}^{K} \pi_k \frac{1}{(2\pi)^{n/2} P^{1/2}} \exp\left(-\frac{1}{2} (Y_{n,k})^T P Y_{n,k}\right) \prod_{i=1}^{D} \frac{Z_{n,ki}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} (Y_{n,ki})^2\right)\right)
\]

**Theorem**

*Under suitable conditions, the likelihood function is bounded above in bounded region; non-decreasing and negative semi-definite w.r.t density \(Z_{n,ki}\); may contain both local minimum and local maximum w.r.t transformed variables \(Y_{n,k}\).*

- Nondecreasing nature of likelihood function in EM algorithms. Check and monitor change to likelihood function during the iteration process (Wu (1983) [19]);
- Global convergence to local maximum.
- Model selection is conducted through AIC information criteria.
Tests on Empirical Road Data

- Comparison to empirical path travel time distribution for path AB-BC-CD

Table: p-value of two-sample K-S test compared with the empirical distribution

<table>
<thead>
<tr>
<th></th>
<th>GMM</th>
<th>GCMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>18</td>
<td>46</td>
</tr>
</tbody>
</table>

- Estimated distributions when AIC indicates three clusters for both:

Figure: Comparison of pdf (Red: GMM; Blue: GCMM; Black: Empirical) and QQplot Empirical(x) v.s. GCMM(y)
Simulation Tests for GCMM

- Simulate data with a three-copula GCMM model

**Table:** p-value of two-sample K-S test compared with the simulated distribution

<table>
<thead>
<tr>
<th></th>
<th>GMM</th>
<th>GCMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>0.0002</td>
<td>0.1304</td>
</tr>
</tbody>
</table>

- GCMM achieves better fitting with fewer clusters.

**Figure:** Clusters for GMM v.s. Clusters for GCMM
Alternative decision rules are developed, which include:

- Value-at-Risk;
- Exponential utility function;
- The Mean-Variance rule;
- The Area-Ratio rule;
- Stochastic dominance: First order (FSD)/ Second order (SSD);

Different decision rules may yield different routing suggestions to travelers.
Outline

1 Introduction to Transportation Risk Management

2 Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)

3 Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
   - Modeling Travel Time Changes as Continuous Time ARMA\((p, q)\) Processes
   - Pricing Travel Time Derivatives

4 Conclusion and Future Work
Travel Time Derivatives

- How could travelers hedge economic risk due to traffic delays?
- A travel time derivative contract is a financial instrument whose price is derived from the values of travel time measurements.
  - Hedge economic risk given routing patterns;
  - Further change routing patterns to minimize economic risk.

Travel time call option:
If the experienced travel time on the link is higher than 4 minutes, the payoff is $20(T_t - 4)$

Routing choice

Call option price $5$

Call option price $6$

More compensation against congestion

Call option payoff more $20(T_t - 4)$

Call option is more valuable
Comparison with Traditional Road Tolls

Comparison of the payoffs:

<table>
<thead>
<tr>
<th>Traffic Condition</th>
<th>Traffic payoff</th>
<th>Toll payoff</th>
<th>Derivative payoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good traffic</td>
<td>$T_{good}$</td>
<td>$-P$</td>
<td>$-P(T)$</td>
</tr>
<tr>
<td>Bad traffic</td>
<td>$T_{bad}$</td>
<td>$-P$</td>
<td>$\alpha(T_{bad} - K) - P(T)$</td>
</tr>
</tbody>
</table>

An example of a US Congestion Index Futures contract:

<table>
<thead>
<tr>
<th>Contract Size</th>
<th>US $20 times the respective CME USA Congestion Index</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Index Product Description</th>
<th>Congestion levels for U.S. Cities</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Measurement definition</th>
<th>The travel times based on specific routes in two particular cities are reported and averaged: 1: Along Fifth Avenue between 59th st and Washington Square, New York 2: Along Michigan Avenue between South Lake Drive and Roosevelt Road, Chicago</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Pricing Unit</th>
<th>US Dollars (US $) per index point</th>
</tr>
</thead>
</table>

| Tick Size | 1 index point (= US $20 per contract) |
Market Making of Travel Time Derivatives

- Market participants on both sides:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Hedging Motivation</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual travelers</td>
<td>Traffic delay, business delay</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>extra charges due to bad QOS</td>
<td></td>
</tr>
<tr>
<td>Cargo transportation</td>
<td>Traffic delay due to bad QOS</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>QOS</td>
<td></td>
</tr>
<tr>
<td>Alternative transportation (train)</td>
<td>Less business due to good QOS</td>
<td>H</td>
</tr>
<tr>
<td>Owners of toll roads</td>
<td>Low profit due to good QOS on toll free roads</td>
<td>H</td>
</tr>
<tr>
<td>Gas company</td>
<td>Low overall gas consumption</td>
<td>H</td>
</tr>
<tr>
<td>Banks</td>
<td>Market Making</td>
<td>B/H</td>
</tr>
<tr>
<td>Portfolio managers</td>
<td>Risk diversification Speculation</td>
<td>B/H</td>
</tr>
</tbody>
</table>

- Economic impact: There were $63.2 billion in actual congestion costs in the 85 urban areas in the U.S. in 2002. It is estimated that public transportation saved an additional $20 billion in congestion costs for this group.
Outline

1. Introduction to Transportation Risk Management

2. Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)

3. Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
     - Modeling Travel Time Changes as Continuous Time ARMA$(p, q)$ Processes
     - Pricing Travel Time Derivatives

4. Conclusion and Future Work
Travel time derivatives should be based on:

- Periodical travel time measurements (loop detectors);
- Average travel time on a given path or several paths;
- Indexed by arrival time for calculating payoff;
- Predefined measurement method.

**Definition**

A standard measurement of travel time is the average travel time reported from specific travel time data providers on a path at a given time;
Products based on instantaneous travel time measurements.

**Definition**

Call option on a certain travel time. Consider the link \( l \) a specific time instant \( t \) in the future. If the travel time shown by the standard measurement at \( t \) (denoted as \( T \)) is higher than a given \( K \), then there is a payment \( \alpha(T - K) \) to the option buyer; if lower, there is no payment. \( \alpha \) is the leverage coefficient.

Products based on long term travel time measurements.

**Definition**

Cumulative travel time index (CTT): The CTT index over a time interval \( [t_1, t_2] \) is defined as the sum of the daily standard measurement plan in a given time period.

\[
CTT(t_1, t_2) = \sum_{i=1}^{n} T_{t_i} I_{t_i \in [t_1, t_2]}
\]
Outline

1. Introduction to Transportation Risk Management
2. Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)
3. Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
   - Modeling Travel Time Changes as Continuous Time ARMA($p, q$) Processes
   - Pricing Travel Time Derivatives
4. Conclusion and Future Work
The Model for Travel Time

- Given a probability space \((\Omega, \mathcal{F}, \mathcal{F}_t, P)\), the model of travel time is:
  \[
  T_t = r_t + s_t + Y_t, \quad t = 0, 1, 2, \ldots
  \] \hspace{1cm} (11)

- The trend part is a linear function over time
  \[
  r_t = a + bt
  \]

- The seasonal part contains:
  \[
  s_t = b_t + w_t
  \]

- The daily part:
  \[
  b_t = k + \sum_{1}^{I_s} a_t (\sin 2i\pi(t - f_i)/T_d) + \sum_{1}^{J_s} b_t (\cos 2i\pi(t - g_i)/T_d)
  \]

- The weekly part:
  \[
  w_t = k + \sum_{1}^{I_s} a_t (\sin 2i\pi(t - f_i)/T_w) + \sum_{1}^{J_s} b_t (\cos 2i\pi(t - g_i)/T_w)
  \]

- \(Y_t\) is the driving process. Define \(C_t = dY_t\)
Remove Trend and Seasonality from Travel Time Data

Figure: Trend v.s. Weekly Pattern

Figure: Daily Pattern v.s. Driving process (residuals)
Continuous Time ARMA \((p, q)\) Model

- Definition of CARMA\((p, q)\) (Brockwell (2001) [3]):
  \[
a(D)C_t = b(D)B_t
\]
  with coefficients \(a_1, a_2, \ldots, a_p\) and \(b_0, b_1, \ldots, b_q\) for \(p > q\)
  \[
a(z) = a_0 z^p + a_1 z^{p-1} + \cdots + a_p
\]
  \[
b(z) = b_0 + b_1 z + \cdots + b_q z^q
\]
  The operator \(D\) denotes differentiation with respect to \(t\), which is in the formal sense for the Brownian Motion.
- The state space representation is:
  \[
  C_t = B X_t
  \]
  and
  \[
  dX_t = AX_t dt + e dB_t
  \]
  with
  \[
  B = [b_0, b_1, \ldots, b_q, 0, \ldots, 0]
  \]
  \[
  X_t = (X_{t,0}; \ldots; X_{t,p-1}) \text{ and the first } p-1 \text{ element of } X_t \text{ is defined as}
  \]
  \[
  X_{t,j} - X_{0,j} = \int_0^t X_{u,j+1} du
  \]
Discussion around $p$ and $q$

- Parameters for CARMA($p, q$)
  \[
  A = \begin{bmatrix}
  0 & 1 & 0 & \cdots & 0 \\
  0 & 0 & 1 & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \vdots & \ddots & 1 \\
  -a_p & -a_{p-1} & -a_{p-2} & \cdots & a_1
  \end{bmatrix}
  \]
  \[
e = [0, 0, 0, 0, 0, 1]^T
  \]

- The spectral density for a CARMA($p, q$) process is defined as
  \[
f(\omega) = \frac{\sigma^2 \beta(i\omega)\beta(-i\omega)}{2\pi\alpha(i\omega)\alpha(-i\omega)}
  \]

- The condition of stationary CARMA processes requires the roots of the polynomial $\alpha(z)$ to have negative real-parts and that $p > q$, (Tomasson (2011) [17]).

- When $p \leq q$, CARMA ($p, q$) has recently been extended into generalized random processes (Brockwell (2010) [2]). However, the processes under such settings are not well defined in the usual sense.
Consider the CARMA\((p, q)\) model as a hidden state model
- Maximum likelihood estimation
- Kalman filter extracts hidden state;
- Parameters are estimated based on output and observations.
- The initial values and orders are set via first order approximation.
- Discrete time solution of a CARMA \((p, q)\) process

\[
x_{t+\delta t} = e^{A\delta t}x_t + \int_t^{t+\delta t} e^{t+\delta t-u}e^p\sigma_u dB_u
\]

The noise term is of variance \(\int_0^{\delta t} e^{A\sigma u}e^p e^{At}u du\).
- Initial values are obtained by fitting an ARIMA\((p, d, q)\) model.

**Final estimation:**

<table>
<thead>
<tr>
<th>Path</th>
<th>Parameter</th>
<th>ARMA</th>
<th>Initial Values</th>
<th>Final Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cummings Skwy -</td>
<td>AR-1</td>
<td>-0.03</td>
<td>0.2065</td>
<td>0.2512</td>
</tr>
<tr>
<td>Maritime Academy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berkeley-Davis</td>
<td>AR-1</td>
<td>0.96</td>
<td>0.2071</td>
<td>0.1780</td>
</tr>
<tr>
<td></td>
<td>AR-2</td>
<td>-0.13</td>
<td>0.0326</td>
<td>0.0685</td>
</tr>
<tr>
<td></td>
<td>MA-1</td>
<td>-0.66</td>
<td>0.1356</td>
<td>-0.6296</td>
</tr>
</tbody>
</table>
Outline

1. Introduction to Transportation Risk Management

2. Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)

3. Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
   - Modeling Travel Time Changes as Continuous Time ARMA($p, q$) Processes
   - Pricing Travel Time Derivatives

4. Conclusion and Future Work
Determination of the Pricing Measure

- Non-unique measure change (Benth (2007) [1]):

\[ Z^\omega(t) = \exp\left(\int \frac{\omega_s}{\sigma_s} dW_s - \frac{1}{2} \int \frac{\omega_s^2}{\sigma_s^2} ds\right) \]

is the density process of the probability measure \( Q \). Under \( Q \), the process

\[ dB_t = dW_t - \frac{\omega(s)}{\sigma(s)} dt = dW_t - \theta_t \]

is a Brownian motion and \( \theta_t \) is called the market value of risk. Since the underlying asset is not tradable, there is no unique risk neutral measure.

- Identify the measure through calibration:
  - Calibrate \( \theta_t \) by a set of HCD futures contracts based on the common underlying travel time series by minimizing the mean squared difference between modeled price and traded price below (Hamisultane (2007) [10]):

\[ \theta_t = \arg\min\left(\sum_i \left[ (F_{i,\text{market}} - F_{i,HCD}(t_1, t_2, T, t, \theta_t))^2 \right) \right) \]

  - Hedge with other derivative contracts or stocks whose prices are correlated to the performance of transportation system (Carmona (2003) [4]).
A simple call option:

\[ F_T(t_1, t_2, T, t) = e^{-rt} E_Q \{ \max(T_t - T, 0) | \mathcal{F}_t \} \]

A CTT futures contract:

\[ F_{CTT}(t_1, t_2, t) = E_Q \left\{ \int_{t_1}^{t_2} T_t dt | \mathcal{F}_t \right\} \]

A CTT call option contract:

\[ C_{CTT}(t_1, t_2, t, K) = e^{-rt} E_Q \left\{ \max(\int_{t_1}^{t_2} T_s ds - K, 0) | \mathcal{F}_t \right\} \]
Pricing P.D.E

Theorem

If the asset price process is subject to $Y_t$, where $Y_t = \int_0^t X_u du$ and $X_t$ is a CARMA($p, q$) process with $p > q$, the Asian type call option based on it can be priced via the following P.D.E:

$$
v_t + \frac{1}{2}v_{x_{p-1}x_{p-1}} B(t, X^0_t, \ldots, X^{p-1}_t)^2 + v_{x_{p-1}} A^*(t, X^0_t, \ldots, X^{p-1}_t) + \sum_{i=0}^{p-2} x_{i+1} v_{x_i} + \sum_{i=0}^q b_i x_{q-i} v_y + y v_z = rv
$$

where $v = v(t, x_0, \ldots, x_{p-1}, y, z)$ with boundary conditions

$$
limit_{z \to -\infty} v(t, x_0, \ldots, x_{p-1}, y, z) = 0, 0 \leq t < T, x_0 \in \mathbb{R}, \ldots, x_{p-1} \in \mathbb{R}, y \in \mathbb{R}
$$

$$
v(T, x_0, \ldots, x_{p-1}, y, z) = \left( \frac{z}{T} - K \right)^+, x_0 \geq 0, \ldots, x_{p-1} \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}
$$

$$
v(t, 0, \ldots, 0, 0, z) = e^{-r(T-t)} \left( \frac{z}{T} - K \right)^+, 0 \leq t < T, z \in \mathbb{R}
$$
Sketch of Proof for PDE

Construct a multi-dimensional Markov process

- Price option on integrated process $Z_t$ as an Asian type option for $Y_t$;
- Employ the state space representation of CARMA($p, q$) processes.

Consider the following set of S.D.Es

$$
egin{align*}
    dZ_t &= Y_t dt \\
    dY_t &= \sum_{i=0}^{q} b_i X_t^{q-i} dt \\
    dX_{t}^{0} &= X_{t}^{1} dt \\
    \ldots \\
    dX_{t}^{p-2} &= X_{t}^{p-1} dt \\
    dX_{t}^{p-1} &= A^*(t, X_{t}^{0}, \ldots, X_{t}^{p-1}) dt + \sigma dW_t
\end{align*}
$$

The value of the derivative contract is $de^{-r_t}v(t, x_{p-1}, \cdots, x_0, y, z)$. Apply Ito Lemma and derive the P.D.E based on martingale conditions.
Outline

1. Introduction to Transportation Risk Management

2. Estimation of Travel Time Distribution
   - Copula-based Travel Time Estimation
   - Gaussian Copula Mixture Models (GCMM)

3. Financial Derivatives based on Travel Time
   - Intuition and Market Making
   - Product Design
   - Modeling Travel Time Changes as Continuous Time ARMA($p, q$) Processes
   - Pricing Travel Time Derivatives

4. Conclusion and Future Work
Conclusion and Future Work

Contributions of the thesis

- Developed copula methods, more specifically GCMM to estimate travel time related risk;
- Introduced a routing decision framework with alternative decision rules;
- Introduced the financial derivative based on travel time to hedge economic risk.

Future work:

- The properties of GCMM can be studied further;
- Other statistical models can be used to describe the complex dependent structure between travel times;
- More complex stochastic processes can be used to price travel time derivatives;
- Due-equilibrium of the financial system and transportation system can be studied.
Thank you for your attention

• I analyzed the performance of the transportation system using innovative financial mathematic methods;

• I employed the financial derivative based on performance measures of the transportation system;

• I shared the beauty of Operations Research and Financial Engineering!
A1: Transportation Risk Measures

- How do we measure the risk given a travel time distribution?
- Definition of a transportation risk measure:

**Definition**

A transportation risk measure \( \rho \) is the calculation determining the amount of time to be reserved for possible traffic delays. It is a mapping from a set of random variables to the real numbers, which satisfies the following properties:

- \( \forall a \in R, \rho(a + X) = \rho(X) + f(a) \): When a deterministic delay is added to a stochastic travel time, the risk introduced by this change is a positive increasing function of \( a \) as well.
- If \( X_1 > X_2 \) then \( \rho(X_1) > \rho(X_2) \): if travel time on a link is constantly larger than the travel time on another, then the corresponding risk measure is larger as well.

- A traveler calculates transportation risk measures for available candidate paths at each decision point.
- Value-at-risk, exponential utility function and etc.
A1: Area Ratio Rule

**Definition**

(AR): If the area ratio of the CDF curves of two link travel time $X$ and $Y$ before some critical point $t$ is equal to or greater than a threshold $\epsilon$, $X$ is dominating $Y$; otherwise, $Y$ is dominating $X$. That is

$$X \succeq Y \iff_{AR} \frac{\int_0^t F_A(x)dx}{\int_0^t F_B(x)dx} \geq \epsilon$$

(12)

**Theorem**

*Suppose travelers enter a path according to a poisson process $\lambda$, the cumulative distribution function of travel time is $F$, then according to the properties of poisson process, $P(M_t = k) \sim \text{Poisson}(\lambda')$, where $M_t$ is the number of travelers who complete the route by time $t$. Here, $\lambda' = \lambda \ast \frac{1}{t} \int_0^t F(s)ds$*
A1: Other Decision Rules

- Mean-Variance, Area Ratio, Stochastic Dominance: First order (FSD)/ Second order (SSD) and etc.

**Definition**

Area Ratio (AR): If the area ratio of the CDF curves of two link travel time $X$ and $Y$ before some critical point $t$ is equal to or greater than a threshold $\epsilon$, $X$ is dominating $Y$; otherwise, $Y$ is dominating $X$. That is

$$X \succeq Y \Leftrightarrow_{AR} \frac{\int_0^t F_A(x)dx}{\int_0^t F_B(x)dx} \geq \epsilon$$  \hspace{1cm} (13)

**Theorem**

Suppose travelers enter a path according to a poisson process $\lambda$, the cumulative distribution function of travel time is $F$, then according to the properties of poisson process, $P(M_t = k) \sim \text{Poisson}(\lambda')$, where $M_t$ is the number of travelers who complete the route by time $t$. Here, $\lambda' = \lambda \ast \frac{1}{t} \int_0^t F(s)ds$

- The ratio of the expected number of travelers who complete the route by time $t_l$ under the two given travel time distributions.
A1: Stochastic Dominance

Definition

First-Order Stochastic Domination (FSD): The random variable $X$ first-order stochastically dominates the random variable $Y$, if $P(X > a) \geq P(Y > a), \forall a$.

- If a candidate link is dominated by other candidate links in terms of FSD, then that link is preferred by travelers when making routing choices.

Definition

Second-Order Stochastic Domination (SSD): Suppose the random variables $X$ and $Y$ have support on $[l; u]$. Then $X$ second-order stochastically dominates $Y$ if

$$\int_{l}^{a} P(X > t)dt \geq \int_{l}^{a} P(Y > t)dt, \forall a \in [l, u]$$

- If a candidate link is dominated by other candidate links in terms of SSD, then that link is preferred by travelers when making routing choices.
Decision statistics based on different risk measures (in seconds)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$(\mu, \sigma)$</th>
<th>Exponential utility $a=-1/250$</th>
<th>VAR(Quantile) 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>(258.9, 48.07)</td>
<td>0.3850</td>
<td>286.6</td>
</tr>
<tr>
<td>AFB</td>
<td>(245.4, 98.12)</td>
<td>0.3808</td>
<td>321.6</td>
</tr>
<tr>
<td>Preference</td>
<td>None</td>
<td>AB</td>
<td>AB</td>
</tr>
<tr>
<td>Parameter</td>
<td>AR</td>
<td>FSD</td>
<td>SSD</td>
</tr>
<tr>
<td>$t_U=1500$</td>
<td></td>
<td>first violation</td>
<td>first violation</td>
</tr>
<tr>
<td>AB</td>
<td>1239</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>AFB</td>
<td>1255</td>
<td>259.7</td>
<td>$\infty$</td>
</tr>
<tr>
<td>Preference</td>
<td>AFB</td>
<td>None</td>
<td>AB</td>
</tr>
</tbody>
</table>

Different risk measures lead to different routing decisions, providing more options to travelers.
A2: Sensitivity tests for GCMM on Observable Scenarios

Sensitivity w.r.t Lasso parameter and w.r.t number of scenarios for links AB and BC.

Figure: Left: Change of estimated pdf as the Lasso penalty changes. (Empirical: Red dot; Independent: Yellow; v=0.001: Cyan; v=0.01: Red; v=0.05: Blue; v=0.1: Black; v=0.5: Green), Right: Tail approximation when number of scenarios is 16 based on congestion level.
GCMM employs data more efficiently:

Figure: Comparison of GMM, GCMM base model and GCMM with unsynchronized data
GCMM updates mixtures of copulas and the marginal distributions separately:

Figure: Comparison of GMM, GCMM base model and GCMM with unsynchronized data
A4: Expectation Maximum Algorithm for GCMM 1

- **Expectation Step:**

  \[
  D_{nk}^m = \prod_{i=1}^{D} \frac{Z_{n,ki}^m}{\sqrt{2\pi}} \exp\left(\frac{1}{2} (Y_{n,ki}^m)^2\right) \quad (14)
  \]

  \[
  r_{nk}^m = \frac{\pi_k^m D_{n,k}^m |P_k^m|^{1/2} \exp\left(-\frac{1}{2} (Y_{nk}^m)^T P_k^m, -1 Y_{nk}^m\right)}{\sum_{j=1}^{K} \pi_j^m D_{n,j}^m |P_j^m|^{1/2} \exp\left(-\frac{1}{2} (Y_{nj}^m)^T P_j^m, -1 Y_{nj}^m\right)} \quad (15)
  \]

- **Maximum Step:**

  \[
  \pi_k^m = \frac{\sum_{n=1}^{N} r_{nk}^m}{N} \quad (16)
  \]

  \[
  P_k^m = \frac{\sum_{n=1}^{N} r_{nk}^m Y_{nk}^m (Y_{nk}^m)^T}{\sum_{n=1}^{N} r_{nk}^m} \quad (17)
  \]
Varying-marginal-distribution GCMM (Base Model): In each iteration, update the marginal link travel time probability distribution functions $F_n$ according to the posterior classification probability $r_{nk}$.

\[
F_{ki}^m(y) = \frac{\sum_n r_{nk}^m 1_{x_{ni} \leq y}}{\sum_n r_{nk}^m} \tag{18}
\]

\forall k\text{-th copula, } i\text{-th dimension (kernel smoothing can be applied)}

Varying-marginal-distribution GCMM with Extra Data (Extra-Data):

- In Expectation step:
  - update $r_{nk}^m$ for synchronized data;
  - update $r_{ni,k}'^m$ for un-synchronized data using the following Bayes formula:

\[
r_{ni,k}'^m = \frac{\pi_k^m f_{ki}(x_{ni})}{\sum_{j=1}^K \pi_j^m f_{ki}(x_{ni})} \tag{19}
\]

- In each iteration, update the marginal link travel time cdfs $F_n$ according to $r_{nk}$ and $r_{ni,k}'$. \forall k\text{-th copula, } i\text{-th dimension}:

\[
F_{ki}^m(y) = \frac{\sum_n r_{nk}^m 1_{x_{ni} \leq y} + \sum_{ni} r_{ni,k}'^m 1_{x_{ni} \leq y}}{\sum_n r_{nk}^m + \sum_{ni} r_{ni,k}'^m} \tag{20}
\]
A5: Tests on Empirical Road Data for GCMM Extra Data

- Comparison to empirical path travel time distribution for path AB-BC-CD

**Table:** p-value of two-sample K-S test compared with the empirical distribution

<table>
<thead>
<tr>
<th></th>
<th>GMM</th>
<th>Base Model</th>
<th>Extra-Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>0.05</td>
<td>0.96</td>
<td>0.11</td>
</tr>
</tbody>
</table>

- Estimated distributions when AIC indicates three clusters for both:

Figure: Comparison of pdf (Red: GMM; Blue: Base Model; Cyan: Extra Data; Black: Empirical) and QQplot Empirical(x) v.s. Extra Data(y)
A5: Simulation Tests for GCMM

- Simulate data with a three-copula GCMM model

**Table:** p-value of two-sample K-S test compared with the simulated distribution

<table>
<thead>
<tr>
<th></th>
<th>GMM</th>
<th>Base Model</th>
<th>Extra-Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>0.0002</td>
<td>0.1304</td>
<td>0.1003</td>
</tr>
</tbody>
</table>

- Comparison of clusters.

**Figure:** Clusters for GMM v.s. Clusters for Extra-Data
A5: QQ-plots of Simulated(x) v.s. Estimated(y) CDFs

Figure: Comparison of pdf (Red: GMM; Cyan: Base Model; Blue: Extra Data; Black: Simulated) and QQ plots for GMM

Figure: QQ plots for Base Model and for Extra-Data
A6: Global Properties of the GCMM Likelihood Function

- Log likelihood function:

$$L = \sum_{n=1}^{N} \ln\left(\sum_{k=1}^{K} \pi_k \frac{1}{(2\pi)^{n/2} P^{1/2}} \exp\left(-\frac{1}{2}(Y_{n,k})^T P Y_{n,k}\right)\right)$$

- Derivatives w.r.t marginal density and transformed variables.

$$\frac{\partial L}{\partial Z_{n,k,j}} = \frac{1}{S_n} \pi_k \frac{1}{(2\pi)^{n/2} P^{1/2}} \exp\left(-\frac{1}{2}Y_{n,k}^T P Y_{n,k}\right) \prod_{i \neq j} \frac{Z_{n,ki}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}Y_{n,ki}^2\right) \geq 0$$

$$\frac{\partial L^2}{\partial Z_{n,k,j}^2} = \frac{1}{S_n^2} (0 - P_{n,k,j}^2) \leq 0$$

$$\frac{\partial L}{\partial Y_{n,k}} = \frac{1}{S_n} B_{n,k} D_{n,k} (-P + I) Y_{n,k}$$

$$\frac{\partial L^2}{\partial Y_{n,k}^2} = \frac{1}{S_n^2} (B_{n,k} D_{n,k} (Y_{n,k} Y_{n,k}^T S_n + S_n - B_{n,k} D_{n,k} Y_{n,k} Y_{n,k}^T)) (-P + I) (-P + I)$$

- High dimensional nonlinear equations can be derived based on the formulas above to compute the parameters.
A7: Properties of Path Travel Time Data Synchronization Schemes

- Assemble the trips of different travelers to build data for a specific path.

- Scenario frequency considering pseudo path travel time observations:

\[
\hat{\pi}_i = \frac{\hat{N}_i}{\sum_j \hat{N}_j} = \frac{N_i + \Delta_i}{\sum_j (N_j + \Delta_j)}
\]  

(22)
A7: Properties of the Data Synchronization Scheme

Theorem

Assume the travelers are homogeneous in their driving patterns, and the time mismatch when constructing pseudo path observations is bounded by $\theta_0$, then the frequency of a path scenario estimated using both original and pseudo path observations converges to the true frequency as the number of data points increases to infinity. $\hat{\pi}_i \rightarrow \pi_i$, as $N \rightarrow \infty$ if one of the following three conditions is satisfied:

- if the weights of pseudo observations exponentially decay, as $N \rightarrow \infty$;

$$\hat{\pi}_i = \frac{N_i + \exp(-n_i/n_0)\Delta_i}{\sum_j (N_j + \exp(-n_j/n_0)\Delta_j)}$$

- if the following set of assumptions hold:
  - the arrival events for different scenarios form an independency (Cinlar (2011)[6]).
  - the entry of a traveler to a given link in a certain scenario is a Poisson process with parameter $\lambda$
  - the behavior of different travelers is independent, including the starting time and the route choice to different links.
Definition and Pricing Method for Congestion Days

Definition:

High Congestion Days (HCD): Let $T_i$ denotes the mean of a standard measurement plan on day $d_i$ and $C$ as a specified reference value. The high congestion-days, $HCD_i$ on that day are defined as $HCD_i = \max(T_i - C, 0)$ Then the HCD for a given time period $[t_1, t_2]$ is defined as

$$HCD(t_1, t_2) = \sum_{i=1}^{n} HCD_i 1_{d_i \in [t_1, t_2]}$$

Pricing:

A HCD futures contract:

$$F_{HCD}(t_1, t_2, T, t) = E_Q \left\{ \int_{t_1}^{t_2} \max(T_t - T, 0) dt | \mathcal{F}_t \right\}$$

A HCD call option contract:

$$C_{HCD}(t_1, t_2, T, t, K) = e^{-rt} E_Q \left\{ \max(\int_{t_1}^{t_2} \max(T_s - T, 0) ds - K, 0) | \mathcal{F}_t \right\}$$
F.E. Benth and J. Benth.  
The volatility of temperature and pricing of weather derivatives.  

Peter J Brockwell and Jan Hannig.  
Carma (p, q) generalized random processes.  

PJ Brockwell.  
Continuous-time arma processes.  

R Carmona and A Danilova.  
Hedging financial instruments written on non-tradable indexes.  
*Unpublished manuscript, 2003.*
Adaptive Kalman filter based freeway travel time estimation.  
In 84th TRB Annual Meeting, Washington DC, 2005.

Erhan Çinlar.  
Probability and stochastics, volume 261.  

DJ Dailey, ZR Wall, SD Maclean, and FW Cathey.  
An algorithm and implementation to predict the arrival of transit vehicles.  

Arthur P Dempster, Nan M Laird, Donald B Rubin, et al.  
Maximum likelihood from incomplete data via the em algorithm.  
J. Friedman, T. Hastie, and R. Tibshirani.  
Sparse inverse covariance estimation with the graphical lasso.  

H. Hamisultane.  
Extracting information from the market to price the weather derivatives.  

Ryan Herring, Aude Hofleitner, Pieter Abbeel, and Alexandre Bayen.  
Estimating arterial traffic conditions using sparse probe data.  

Timothy Hunter, Ryan Herring, Pieter Abbeel, and Alexandre Bayen.  
Path and travel time inference from gps probe vehicle data.  
*NIPS Analyzing Networks and Learning with Graphs*, 2009.
CD Mark, AW Sadek, and D. Rizzo.
Predicting experienced travel time with neural networks: a PARAMICS simulation study.

R.B. Nelsen.
An Introduction to Copulas.

Sangjun Park, Hesham Rakha, and Feng Guo.
Multi-state travel time reliability model: Impact of incidents on travel time reliability.
M Sklar.
Fonctions de répartition à n dimensions et leurs marges.
Université Paris 8, 1959.

Helgi Tómasson.
Some computational aspects of gaussian carma modelling.

H. White.
Estimation, Inference and Specification Analysis.

CF Jeff Wu.
On the convergence properties of the em algorithm.
Lei Xu and Michael I Jordan.
On convergence properties of the em algorithm for gaussian mixtures.
Neural computation, 8(1):129–151, 1996.

Xianyuan Zhan, Samiul Hasan, Satish V Ukkusuri, and Camille Kamga.
Urban link travel time estimation using large-scale taxi data with partial information.