SAFETY EXPERIENCE OF LARGE TRUCKS
AN ANALYSIS OF SAMPLE SIZE REQUIREMENTS

by

Gordon A. Sparks
Professor of Civil Engineering
Department of Civil Engineering
University of Saskatchewan

Andrew T. Horosko
Director of Technical Research
Saskatchewan Highways and Transportation

Anne Smith
Senior Associate
Decision Focus Incorporated
Los Altos, California

Second International Symposium On
Heavy Vehicle Weights and Dimensions
Kelowna, British Columbia
June 18 - 22, 1989

SAFETY EXPERIENCE OF LARGE TRUCKS
AN ANALYSIS OF SAMPLE SIZE REQUIREMENTS

by

Gordon A.Sparks [1]
Andrew T.Horosko [2]
Anne Smith [3]

ABSTRACT

Considerable debate has emerged in Canada in terms of the relative safety of different types of large trucks (i.e. double trailer units versus single trailer units; A-Train versus B-Train versus C-Train configurations etc.). This paper presents an analysis of the sample size required in order to have a reasonable expectation of detecting statistically significant differences in safety experience for varying levels of actual differences in accident rates. The results indicate that the sample size required to detect a 10-20% difference in (actual) accident rates is in the order of $1-2 \times 10^9$ vehicle kms of travel. Smaller samples could only be expected to detect larger actual differences. These results suggest that:

1) it is very important to include an appropriate statistical analysis as part of accident re-search analysis if differences in the order of 20% are of interest, and

2) the value of much of the research reported on the relative safety experience of various vehicle types may be limited and/or questionable because of limited sample sizes and/or the lack of appropriate statistical analysis.

SAFETY EXPERIENCE OF LARGE TRUCKS
AN ANALYSIS OF SAMPLE SIZE REQUIREMENTS

INTRODUCTION

The debate in Canada related to the use of larger trucks has
increased in recent months. The increased level of concern and
debate is in part related to the release of recommendations
resulting from research undertaken by the Roads and
Transportation Association of Canada (RTAC). RTAC released the
recommendations from a multi-year, multi-million dollar research
program in the spring of 1987. This research was directed at
assessing the benefits and costs associated with permitting
larger and heavier trucks to operate on Canadian highways. [4]

The RTAC research concluded, in essence, that the significant
economic benefits associated with the operation of larger and
heavier trucks outweighed the road and other associated costs and
it was therefore deemed to be in the public interest to allow
larger and heavier trucks to operate on Canadian public highways.
The RTAC recommendations were first considered by the Council of
Ministers Responsible for Transportation and Highway Safety in
September of 1987 and, with minor modifications, were accepted in
February 1988.

Major opposition to the RTAC recommendations to permit larger and
heavier trucks on public roads came from the railways and highway
users represented by the Canadian Automobile Association (CAA).

The opposition put forward by the railways was largely self-
serving in recognition of the potentially significant increase in
competition from trucks made possible by the permitted increases
in truck size and weights. The opposition put forward by the CAA
was based upon a concern for public safety. The general position
put forward was that bigger trucks presented an increased hazard
and inconvenience to the motoring public.

There is little question that safety is a major consideration
relative to highway transport generally, and the use of larger
and heavier vehicles in particular. Unfortunately much of the
debate as well as much of what is published on this subject is
highly emotional, and often lacks a substantive factual basis.
The highly emotional nature of the debate can be attributed in
large part to the significant personal vested interests possessed
by the participants. The lack of objective factual information
on the other hand is attributable to the complexity of the issue
combined with the lack of detailed data and analysis. The

2

situation is further complicated by the probabilistic nature of highway safety.

Accident experience of various types of vehicles are often compared but seldom are these comparisons undertaken in a manner consistent with sound statistical methods. As a result, many of the comparisons are interesting but in a scientific sense are not particularly meaningful. Some are even misleading.

This paper identifies, illustrates and discusses one of the methodological difficulties encountered by researchers attempting to present an unbiased, objective assessment of the facts related to the safety experience of large trucks operating on public highways. In particular, this paper addresses the issue of sample size required to detect statistically significant differences in accident experience between different types of vehicles. The analysis presented would for example be applicable to determining how much data (in terms of vehicle kms of travel) would be required to detect statistically significant differences in accident experience between single versus double trailer units, or between A, B, and C-Train configurations.

ANALYSIS FRAMEWORK

In the context of the above, the issue of interest is whether there is a statistically significant difference between one vehicle type's accident experience relative to that of another vehicle type (i.e. single trailer unit versus double; A-Train configurations versus B-Train etc.). In statistical terms, such questions can be addressed by testing for statistically significant differences between means.

In this case it is easier to develop the details in the context of a specific situation and then generalize the methodology and conclusions. Therefore the following will be developed in the context of double trailer units versus singles. It will be obvious as to how the methodology could be extended to A versus B or A versus C or B versus C-Trains.

Let:

$\overline{X}_D$ = average number of accidents per 100 million kms of travel for doubles from the sample

$\overline{X}_S$ = average number of accidents per 100 million kms of travel for singles from the sample

3

Then, $(\overline{X}_D - \overline{X}_S)$ will be normally distributed

$$N\left(\mu_D - \mu_S ; \sqrt{\frac{S_D^2}{\eta_D} + \frac{S_S^2}{\eta_S}} \right)$$

Where

μ_D = average number of accidents per 100 million kms
of travel for doubles from the population

μ_S = average number of accidents per 100 million kms
of travel for singles from the population

S_D^2 = sample variance for the doubles
S_S^2 = sample variance for the singles
n_D = number of samples of X_D
n_S = number of samples of X_S

If the number of truck accidents per 100 million kms of
travel is (approximately) Poisson distributed ([5]), then

S_D^2 approximately equals μ_D
S_S^2 approximately equals μ_S

Then if the accident rates for doubles is approximately
equal to singles and both are in the order of 150
accidents per 100 million kms of travel ([6]), and n_D ~
n_S (i.e. sample size for doubles and singles are
equal), then

$(\overline{X}_D - \overline{X}_S)$ will be approximately normally
distributed

$$N\left(\mu_D - \mu_S ; \sqrt{\frac{2S^2}{\eta_D}} \right)$$

If n = 10 (x100 million kms of travel for each doubles
and singles ([7]) for a total of 2 x 10 x 100 million kms
of travel in total, or 2 x 10⁹ kms of travel), then

$(\overline{X}_D - \overline{X}_S)$ will be approximately distributed

$$N\left(\mu_D - \mu_S ; .455 \right)$$

But S^2 is approximately equal to 150 or S = 12.25

therefore,

$(\overline{X}_D - \overline{X}_S)$ will be approximately distributed

4

$$N\left(\mu_D - \mu_S \; ; \; 5.51\right)$$

If a 95% confidence interval is required before accepting an observed difference as real, then we would have to obtain an observed difference of sample means of 1.645 x 5.51 = 9.06. [8] A 10% difference in sample means would be approximately 15 (i.e. 10% x 150). The probability of capturing this 10% difference is:

$$Z = \frac{\text{observed difference}}{\text{standard error}}$$

$$Z = \frac{9.06 - 15}{5.51} = \frac{-5.94}{5.51} = -1.08$$

The area under the standard normal curve (F) for Z = -1.08 equals:

$$F \ (Z \ : \ -1.08) = 85\%$$

Therefore the probability of capturing a 10% difference in accident rate with a sample size of n = 10 (10 x 100 million kms for each doubles and singles or a total of 2×10^9 kms of total truck travel) would be 85%.

The above can be repeated for varying values of n and varying levels of actual differences to get the following:

Probability of Detecting A Difference

n	Actual Difference 10%	20%
1	22%	54%
5	62%	98%
10	85%	99%

From the above there is an 85% chance of finding a significant difference with a simple difference of means test if an actual difference of 10% exists when n = 10. If the actual difference is 20%, then there is a 99% chance of finding a significant

difference if such exists. If n is reduced to 5 or 1, then the
chances of detecting a 10% difference are reduced to 62% and 22%
respectively. If the actual difference is 20% the chances are
larger, i.e. 98% and 54% respectively.

If we reduce the significance level required from 95% to 90%, and
keep n = 10, we can increase the probability of observing a
significant difference when the actual difference is 10% to about
93%. However we increase the probability of a Type I error from
5% to 10%. That is, there is a 10% instead of 5% chance that we
will reject the hypothesis H_0 that the accident rates for doubles
is equal to that of singles in favour of the hypothesis H_1 that
the accident rate for doubles is greater than singles, when in
fact the accident rate for doubles is equal to that of singles.
In a policy sense, this would lead to prohibiting doubles when
there is no reason to prohibit them.

Conversely, this says that there is a 15% probability (100% -85%)
of accepting the hypothesis H_0; that they are equal, and
rejecting the hypothesis H_1 that accident rates for doubles is
greater than that for singles when in fact doubles have an
accident rate 10% greater than singles. (This is the probability
of a Type II error.) This in a policy sense may result in
permitting doubles when in fact they experience 10% more
accidents.

If the primary objective is to protect public safety, the second
error is obviously more serious than the first. In any event
public policy makers would probably want to be reasonably sure
(i.e. 95% or greater) that they didn't permit larger vehicles on
the highway (i.e. doubles) if they experience substantially more
accidents (i.e. 10-20% more than singles). In order to carry out
an analysis that would provide the required information with the
required level of confidence, it would appear as though "n", the
sample size, would have to be in the order of 5-10 for a
confidence level on a Type II error of >90% for real difference
of 20%. This translates into a sample size of 5-10 x 100 million
kms of travel for each doubles and singles, or between 1 - 2 x
10^9 vehicle kms in total (half for doubles and half for singles).
If it is assumed that the average vehicle (singles or doubles)
travelled 200,000 kms per year one would have to have a data base
that included 5000 vehicle - years of data. This could be a
fleet of 5000 vehicles for 1 year, 2500 vehicles for 2 years etc.
Put another way, it would require 50 fleets of 100 vehicles each
for one year, or 10 fleets of 100 vehicles for 5 years. In the
context of Canada, there are probably 50 fleets that operate 100
vehicles or more but the real practicality of undertaking such an
experiment is limited when one reflects back upon the required

assumptions. (See note 5) This in essence requires that all things must remain unchanged during the sampling. That is, vehicle types cannot shift, vehicles must operate on the same types of roads, driving skills must remain constant, and operations in the same traffic levels etc. The likelihood of meeting this requirement over 5000 vehicle - years of travel is remote at best. Any deviation from these assumptions would create "noise" in the sample which in turn would increase the sample size required to detect significant differences at any level of actual difference in accident experience.

CONCLUSION

The above results have important consequences in terms of research and analysis related to determining the relative safety of various vehicle types and/or configurations. That is, the sample size required to detect even 20% difference in accident rates in a "clean" sample is beyond the size that could be reasonably assembled in Canada. This suggests that there is little if any chance of detecting these differences in a "dirty" or "noisy" sample (i.e. shifting composition of vehicle types, travelling different amounts on various roads etc.) which is more representative of the real world data that is available.

Further, the above suggests that conclusions drawn from smaller samples (i.e. as per those sometimes reported) are at best questionable unless the difference in accident experience is large (i.e. >>20% difference).

Given that the practicality of undertaking such a study on a fleet specific basis is limited, one might consider undertaking such an experiment for a particular section of road. A busy section of road handles 100,000 vehicles/day of which 30% could be trucks, with half of these singles and half doubles. If one has a 10 km section of (uniform) roadway, one would have to collect data for 10 years, i.e.

$$100,000 \times 30\% \times 10 \text{ km} \times X \text{ days} = 1 \times 10^9 \text{ vehicle km}$$

$$X = \frac{1 \times 10^9}{100,000 \times .3 \times 10} = \frac{1 \times 10^9}{3 \times 10^5} = .33 \times 10^4 \text{ days}$$

$$= \frac{3300 \text{ days}}{365} = 9.04 \text{ years}$$

The above suggests that large sample sizes are required if differences in actual safety experience in the order of 20% are to be detected. This in turn suggests that it is fundamentally important to undertake appropriate statistical analysis when drawing conclusions about the relative safety experience of different vehicle types.

ENDNOTES

1. Professor of Civil Engineering, Department of Civil Engineering, University of Saskatchewan

2. Director of Technical Research, Saskatchewan Highways and Transportation

3. Senior Associate, Decision Focus Incorporated, Los Altos, California

4. <u>Recommended Regulatory Principles for Interprovincial Heavy Vehicle Weights and Dimensions</u>, RTAC Vehicle Weights and Dimension Study, Implementation Planning Subcommittee, September 1987

5. The number of accidents per unit of travel (i.e. per 100 million vehicle kms of travel) would be expected to be a Poisson distributed random variable if:

    1) The random variable is the number of times that an event occurs in a single trial which is defined as a basic unit of time, space, or quantity;

2) The trials which are non-overlapping are independent of each other;

3) The average number of occurances of the random variable in a single trial remains constant from trial to trial.

In the context of the accident experience of large trucks, (1) and (2) above present no problem. However, (3) presents a major challenge because it essentially says that nothing (i.e. types of trucks, roads, drivers, etc.) can change during the sampling.

6. Data has been assembled on a fleet specific basis for a total of 600 million vehicle kms of truck travel in western Canada over 4 years. This data suggests that the average accident rate for over the road highway operations is in the order of 150 accidents per 100 million vehicle kms for both double and single trailer units.

7. One might be tempted to say that with n = 10, n is small in the context of large - sample tests and that the Student's t Distribution would be more appropriate for small n. Care must be used in the interpretation of n_D or n_S. In the context of our formulation n_D and n_S are the number of samples of $\overline{X}_D$ and $\overline{X}_S$ sampled from a very large population. In this way, n = 10 is not synonymous with a small sample (*). The point might be illustrated as follows. We might have defined $\overline{X}_D$, $\overline{X}_S$ etc. as the number of accidents in 10 million or 1 million vehicle kms of travel instead of the number per 100 million vehicle kms as we did. This would have had the effect of increasing n_D and n_S by a factor of 10 or 100. That is, for 2×10^9 kms of travel we would have had n_D and n_S equal to 100 or 1000. In turn $\overline{X}_D$ and $\overline{X}_S$ would have been reduced by a factor of 10 or 100 from 150 accidents per 100 million vehicle kms to 15 accidents per 10 million vehicle kms or 1.5 accidents per million vehicle kms. Clearly $n_D = n_S = 100$ or 1000 would have been large in the context of sample size and presumably the assumption that the difference of means was normally distributed would be appropriate. It is noted however that nothing has really changed in

9

redefining n other than rescaling $\overline{X}_D - \overline{X}_B$, and $S_D{}^2$ and $S_B{}^2$. Their relative values, which are the important issue in the statistical test, remain unchanged. The requirement for n to be large relates to the sample that $\overline{X}_D$ and $\overline{X}_B$ are drawn from and not n_D and n_B per se. Clearly in our case, $\overline{X}_D$ and $\overline{X}_B$ are drawn from very large samples and therefore it is appropriate to use the normal distribution for this statistical test.

*   <u>Mathematical Statistics with Applications</u>,
    Third Edition, N. Mendenhall et al, PWS
    Publishers, 1986

8. The difference of means is normally distributed. The value for Z equals 1.645 for 95% under a one-tailed test. See any standard Statistics text for details.

SESSION 11 – IMPLEMENTING RESEARCH

Chairman: Byron Geuy, Western Highway Institute

Speakers

1.   **U.S. Truck Size and Weight Study**
     A.J. Balek, J.R. Stowers, H. Weinblatt, Federal Highway Administration

2:   **The Canadian Agreement on Vehicle Weights and Dimensions**
     J. Pearson, RTAC

3   **Truck Size and Weight Issues in Australia**
     P. Sweatman, Australian Road Research Board

4.   **The Transportation Research Board's "Turner Proposal" Study**
     J.R. Morris, Transportation Research Board