16198

Techniques for the Statistical Analysis of Observer Data

John G. Bennett
U.S. Army Tank-automotive and Armaments Command

Ground Vehicle Survivability Symposium
28 March 2001
Monterey, California

Summary

- The Experimental Situation
- Analysis Techniques
  - —Fitting Logistic Curves
  - —Fisher Exact Test
- Comparison of Techniques
- Discussion of Detectability Specifications
- Recommendation

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
28 MAR 2001	**N/A**	**-**

4. TITLE AND SUBTITLE	
Techniques for the Statistical Analysis of Observer Data	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Bennett, John G.	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US ARMY TACOM 6501 E 11 Mile Road Warren, MI 48397-5000**	8. PERFORMING ORGANIZATION REPORT NUMBER **16198**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) **TACOM TARDEC**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) **16198**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

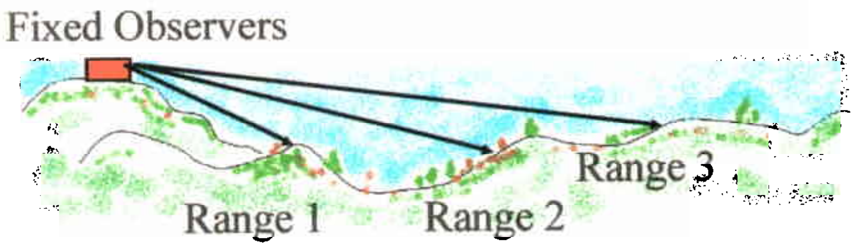
14. ABSTRACT

15. SUBJECT TERMS

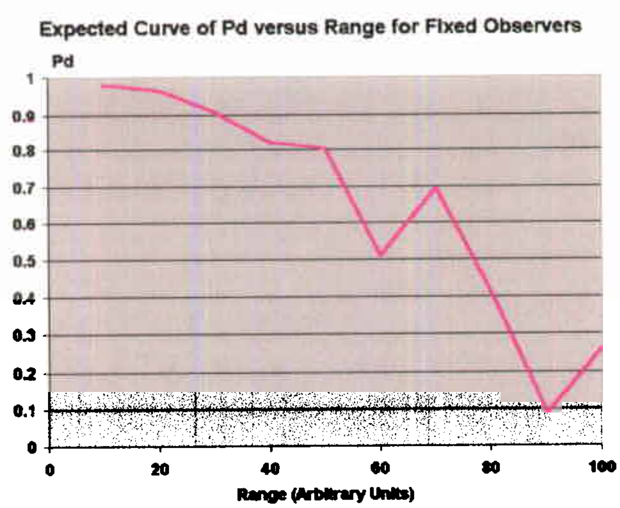
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**15**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

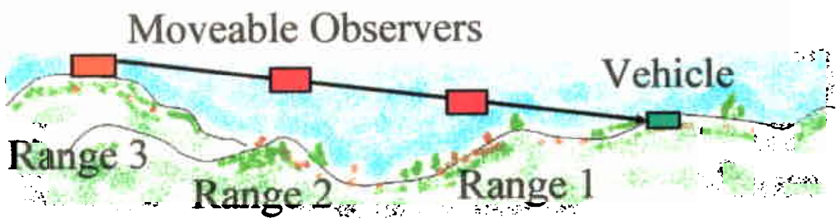
Test with Fixed Observers



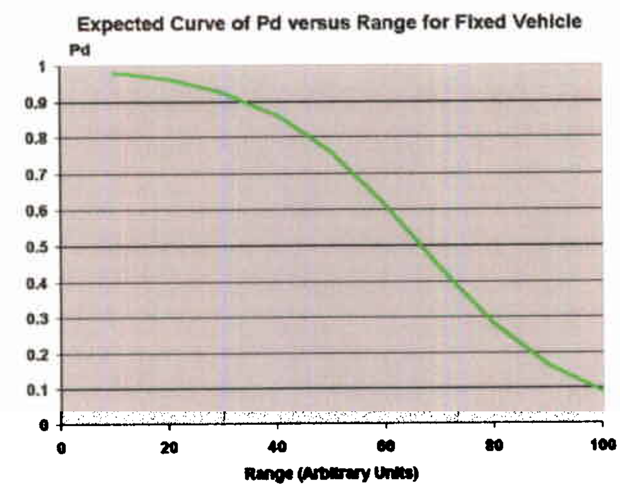
Fixed Observers



Test with Fixed Vehicle



Fixed Vehicle



Analysis Techniques

- Fitting Logistic Curve
- Using Fisher Exact Test

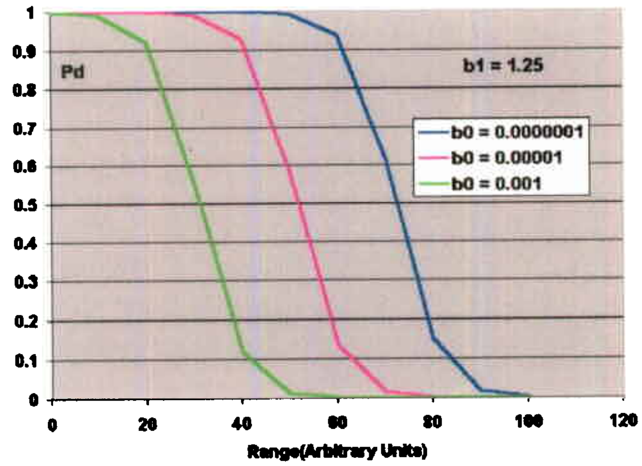
Fitting Logistic Curves

- Based on the logit transformation, logistic curves are defined by

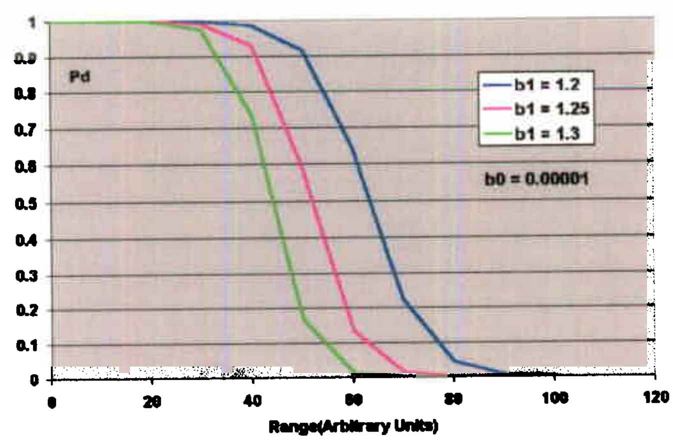
$$P_d = \frac{1}{1 + b_0(b_1^R)}$$

Where Pd is probability of detection, R is range, and the fitting parameters are b_0 and b_1.

Family of Logistic Curves



Family of Logistic Curves



Output fromFitting Curves

- Fitted curves
- Significance of curve parameters, probability that fitted parameters are better than null parameters.
- No value for significance of difference between vehicles.

Fisher Exact Test

	Vehicle A	*Vehicle B*	*Total*
Detect	15	18	33
No Detect	24	15	39
Total	39	33	72
Pd	38%	55%	46%

Null Hypothesis: Vehicle A and Vehicle B have the same Pd.

For 95% confidence, reject null hypothesis if p(Detects<15) > 5%

Fisher Exact Test

Use Hypergeometric Distribution:

	Vehicle A	Vehicle B	Total
Detect	a	b	r_1
No Detect	c	d	r_2
Total	s_1	s_2	N

$$p(a, s_1) = \frac{r_1! r_2! s_1! s_2!}{N! a! b! c! d!}$$

Fisher Exact Test

	Vehicle A	Vehicle B	Total
Detect	15	18	33
No Detect	24	15	39
Total	39	33	72
Pd	38%	55%	46%

$$p(Detects < 15) = p(15) * 0.5 + p(14) + ... + p(0)$$

$$= 9.2\% > 5\%$$

Fisher Exact Test

	Vehicle A	Vehicle B	Total
Detect	15	18	33
No Detect	24	15	39
Total	39	33	72
Pd	38%	55%	46%

We reject the null hypothesis.

We cannot be sure with 95% confidence that Vehicle A is less detectable than Vehicle B.

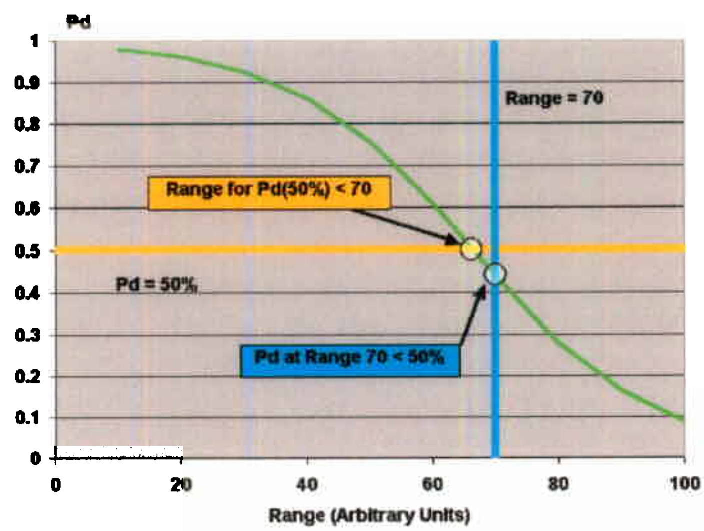
Comparison of Techniques

- Only Fisher Exact Test allows hypothesis testing with the calculation of significance of difference.
- Fitting logistic curves allows interpolation between data points and extrapolation beyond the data.

Detectability Specifications

- **Max Range** for P_d(50%)
- **Max P_d** at Given Range

Max Range and Max P_d Specs



Detectability Specifications

- **Max Range** for $P_d(50\%)$
  - Requires interpolation to calculate a range between measured points.
- **Max P_d** at Given Range
  - Direct calculation at a measured range.

Recommendations

- To compare two vehicles, observation tests should be analyzed by the Fisher Exact Test.
- For a directly testable detectability specification, the maximum P_d at a given range should be required.

Techniques for the Statistical Analysis of Observer Data

14 February 2001

John G. Bennett
U.S. Army Tank-automotive and Armaments Command
Warren, MI 84397-5000

ABSTRACT

For vehicle designers, the main goal of experiments on the observability of combat vehicles is a comparison of the probability of detection for two vehicles as a function of range. This paper addresses the statistical techniques to analyze such data to compare the probability of detection of the two vehicles as a function of range. Two techniques are compared: 1. Fitting logistic curves to the data of the vehicles and 2. Using the Fisher Exact Test to compare the probability of detection of the two vehicles at each range. The paper discusses the issues of background variability and confidence levels for hypothesis testing. Finally, a recommendation is made on how to write a specification for the detectability of a vehicle.

1. The Experimental Situation

Consider Figure 1, a sketch of a typical field test of vehicle detectability. Typically, the experimenter offers observers a large number of views of each vehicle at several ranges and in a variety of backgrounds. The experimenter records the numbers of true, false and missed detections of the target vehicle.

Note that in these experiments with fixed observers range cannot be varied independently of background conditions. Figure 2 shows an experimental setup in which range can be varied independently of the background. This setup, however, requires moving the observers as the target remains fixed. Although such a setup could in principal be employed, costs will generally dictate tests with fixed observers.

Such a moveable observer test should generate a monotonic curve of probability of detection (Pd) versus range, Figure 3, because range has varied independently of the vehicle background. On the other hand, a fixed observer test will generate a Pd versus range curve with background variability superimposed on a monotonic curve, Figure 4.

2. Analysis Techniques

For vehicle designers, the main goal of experiments on the observability of combat vehicles is a comparison of the probability of detection for two vehicles. In this paper, I will compare two techniques for analyzing data from such fixed observer tests:

1. Fitting logistic curves to the data of the vehicles.
2. Using the Fisher Exact Test to compare the probability of detection of the two vehicles at each range.

Logistic Curves

The logistic curve arose from the transformation of probabilities, p, by the so-called logit transformation, namely,

$$y = \log_e \frac{p}{1-p}$$

The logit transformation tends to linearize data in which a proportion is a function of another variable, such as range.

The following equation defines the logistic curve itself:

$$P_d = \frac{1}{1 + b_0(b_1^R)}$$

General statistics software packages, such as SPSS, can fit a logistic curve to data and can calculate a measure of goodness of fit.

Fisher Exact Test

The Fisher Exact Test is designed to compare two proportions to decide if the proportions are significantly different. The test begins with the preparation of a contingency table. For observer tests, a separate table would be prepared for each range.

For example, consider the following table:

	Vehicle A	Vehicle B	Total
Detect	a	b	a+b
No Detect	c	d	c+d
Total	e	f	e+f
Pd	a/e	b/f	(a+b)/(e+f)

The Fisher Exact Test calculates the probability that the observed contingency table or a more extreme table would be the result of random variations of draws from a population with a Pd of (a+b)/(e+f). The Fisher Exact Test uses the hypergeometric distribution to calculate the required probabilities. The hypergeometric distribution describes the probability, for example, of drawing 5 red cards in a 13-card hand from deck of 26 red and 26 black cards.

The Fisher Exact Test supports direct hypothesis testing. A typical null hypothesis would be that Vehicle A and Vehicle B have the same Pd. The experimenter would select a significance level before the test, for example 5%. If the observed contingency table has a probability less than 5%, then the experimenter can state with greater than 95% confidence that the two vehicles have different Pd 's.

3. Discussion

The Fisher Exact Test has advantages over fitting a logistic curve. Because data is compared only at the same range, the effect variability of background with range is avoided. Experimental sample sizes can be calculated for comparison of proportions in order to insure a given difference in Pd will be significant. Moreover, the Fisher Exact Test lends itself to quantitative hypothesis testing.

4. Recommendation on Writing Specifications

The analysis of observer data has a bearing on how specifications should be written. To use quantitative hypothesis tests to judge compliance to specifications, the maximum allowed Pd at given range should be specified instead of the maximum allowed range at Pd of 50%. Consider the following specifications: 1. "The range for a Pd of 50% must be less than R." and 2. "The Pd at a range of R must be less than 50%." The two specifications are equally stringent, but the second specification allows a direct test with a quantitative significance level. On the other hand, the only way to determine the range for Pd of 50% is through interpolation by curve fitting.

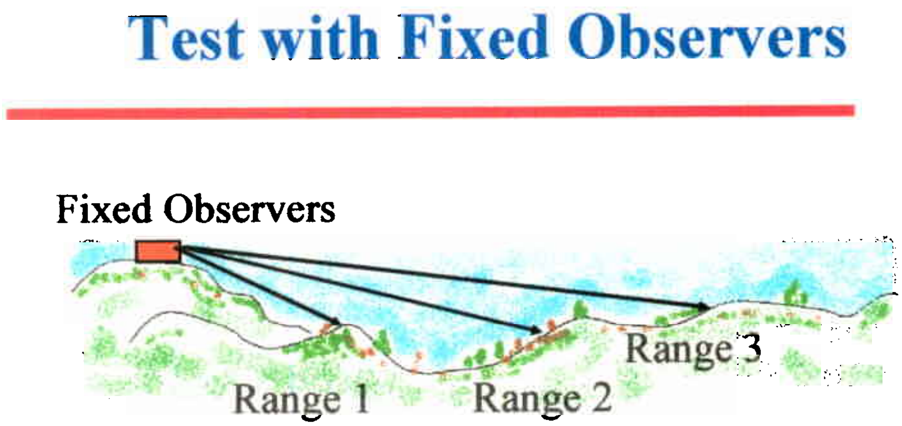


Figure 1. Sketch of an observer test using fixed observers.

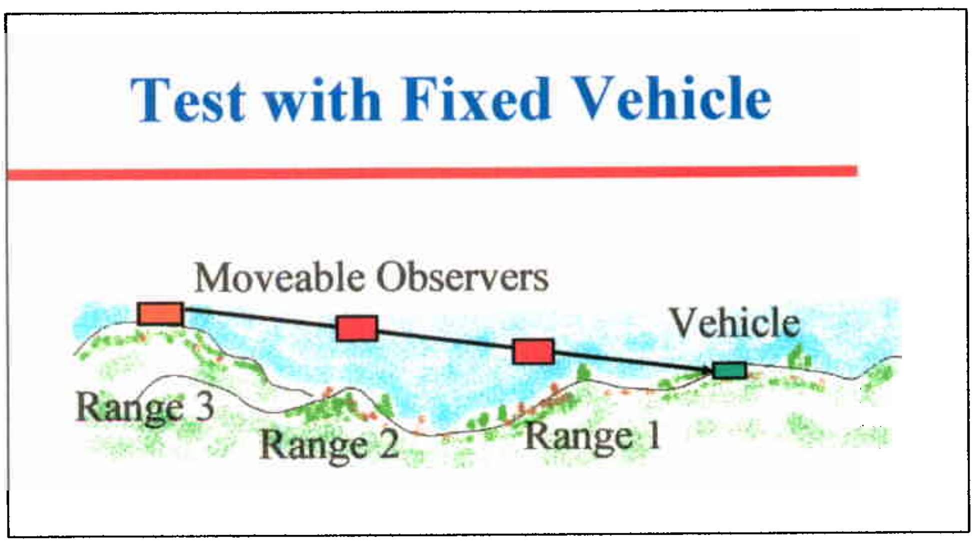


Figure 2. Sketch of an observer test with a fixed target vehicle.

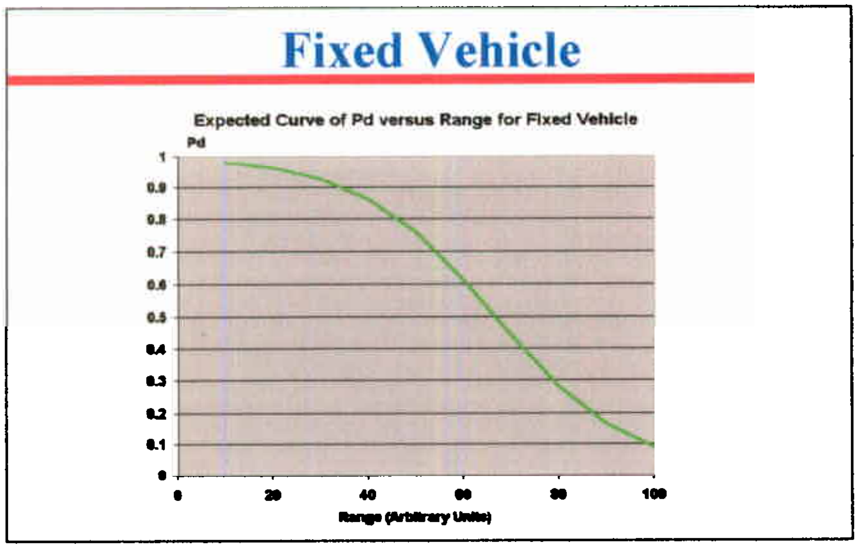


Figure 3. Curve to be expected from a test using a fixed target vehicle.

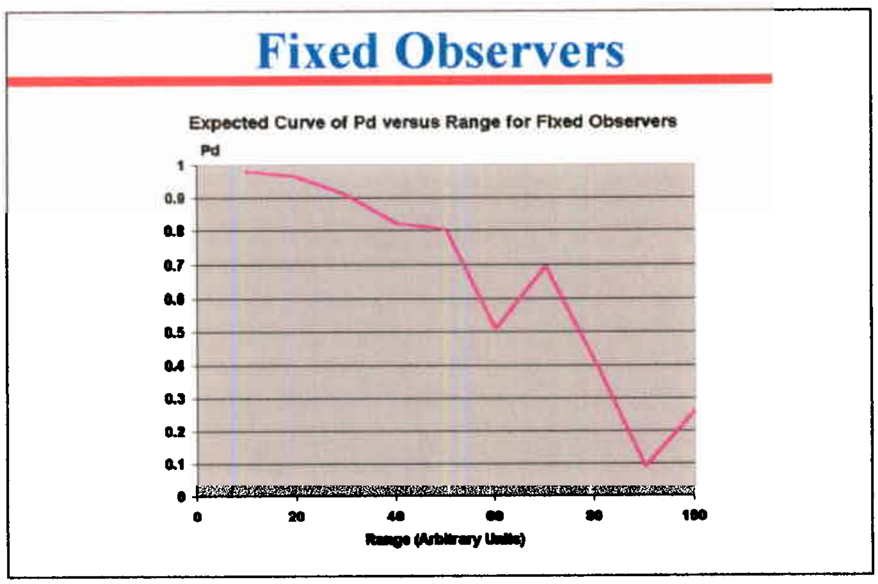


Figure 4. Curve to be expected from an observer test using fixed observers.