412TW-PA-12811



Confidence Intervals for Binary Responses-R50 & the Logistic Model

ARNON M. HURWITZ

AIR FORCE TEST CENTER EDWARDS AFB, CA

OCTOBER, 2012

Approved for public release A: distribution is unlimited.

AIR FORCE TEST CENTER EDWARDS AIR FORCE BASE, CALIFORNIA AIR FORCE MATERIEL COMMAND UNITED STATES AIR FORCE

REPORT DOCUMENTATION PAGE					Form Approved						
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction					CIMB NO. 0704-0188						
data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing											
4302. Respondents should be	aware that notwithstanding an	y other provision of law, no perso	in shall be subject to any penalty	for failing to comply wit	h a collection of information if it does not display a currently						
valid OMB control number. PL		IR FORM TO THE ABOVE ADDI	RESS.	3	DATES COVERED (From - To)						
October	2013	Presentation		5.1	1 October 2013 – 30-October 2013						
		resentation		5a.	CONTRACT NUMBER						
4. IIILE AND	SUBIILE										
				5b.	GRANT NUMBER						
Title: Confidence	e Intervals for Bir	nary Responses									
Subtitles D50 & t	ha Lagistia Mad	2 I		5c.	PROGRAM ELEMENT NUMBER						
Subline: K50 & l	the Logistic Mou	el									
6 AUTHOR(S)				5d.	PROJECT NUMBER						
а ълтт [.]				5e.	TASK NUMBER						
Arnon M. Hurwi	tz (US Air Force))									
				5f.	WORK UNIT NUMBER						
7. PERFORMING ORG	ANIZATION NAME(S)	AND ADDRESS(ES) AN	ND ADDRESS(ES)	8. 1	PERFORMING ORGANIZATION REPORT						
					NUMBER						
Air Force Flig	ht Test Cente	r			412TW-PA-12811						
412 Test Wing											
Edwards AFB CA	A 93524										
9. SPONSORING / MO	NITORING AGENCY	NAME(S) AND ADDRES	S(ES)	10.	SPONSOR/MONITOR'S ACRONYM(S)						
					N/A						
		CE TEST CENT	ER								
	FDW	ARDS AFR CA		11.	SPONSOR/MONITOR'S REPORT						
					NUMBER(S)						
12. DISTRIBUTION / A	VAILABILITY STATE	MENT		•							
Approved for public	release A: distributi	on is unlimited.									
13. SUPPLEMENTAR	YNOTES										
13. SUPPLEMENTAR CA: Air Force Test	Y NOTES Center Edwards AFE	3 CA CC:	012100								
13. SUPPLEMENTAR CA: Air Force Test (Y NOTES Center Edwards AFE	3 CA CC:	012100								
13. SUPPLEMENTAR CA: Air Force Test (Y NOTES Center Edwards AFE	3 CA CC:	012100								
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is	Y NOTES Center Edwards AFE	3 CA CC:	012100	For example y	= {success_failure} for blin-scan radar						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res	Y NOTES Center Edwards AFE a non-linear metho	3 CA CC: d for modeling a bina	012100 ry response variable.	For example, y	= {success, failure} for blip-scan radar						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res	Y NOTES Center Edwards AFE s a non-linear metho ponses cannot be m	3 CA CC: d for modeling a bina odeled using regular l	012100 ry response variable. inear regression. In o	For example, y ur work, many a	= {success, failure} for blip-scan radar applications of logistic regression						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves.	Y NOTES Center Edwards AFE s a non-linear metho ponses cannot be m In the present discu	3 CA CC: d for modeling a bina odeled using regular I ussion, models allowir	012100 ry response variable. inear regression. In o ng independent slope	For example, y ur work, many a s and independ	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple	r NOTES Center Edwards AFE s a non-linear metho ponses cannot be m In the present discu groups of measures.	3 CA CC: d for modeling a bina odeled using regular I ussion, models allowir The question that we	012100 ry response variable. inear regression. In o ng independent slope e consider here is the	For example, y ur work, many a s and independ construction of	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad	Y NOTES Center Edwards AFE s a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50)	3 CA CC: d for modeling a bina odeled using regular I ussion, models allowir The question that we values for two logistic	012100 ry response variable. inear regression. In o ng independent slope consider here is the curves with each valu	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50)	3 CA CC: d for modeling a bina odeled using regular I ussion, models allowir The question that we values for two logistic ieves 50% detection p	012100 ry response variable. inear regression. In o ng independent slope consider here is the curves with each valu probability. This proble	For example, y ur work, many a s and independ construction of ue (viz. R1, R0) a em is the same	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %')	3 CA CC: d for modeling a bina odeled using regular I ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien	012100 ry response variable. inear regression. In o ng independent slope consider here is the curves with each valu robability. This proble	For example, y ur work, many a s and independ construction of ue (viz. R1, R0) a em is the same problem analyt	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction'	d for modeling a bina odeled using regular l ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method	012100 ry response variable. inear regression. In o ng independent slope consider here is the curves with each valu robability. This proble ice. We approach the is. Our results are bas	For example, y ur work, many a s and independ construction of ue (viz. R1, R0) a em is the same problem analyted on the large	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on re	d for modeling a bina odeled using regular l ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the lea	012100 ry response variable. inear regression. In o ng independent slope e consider here is the curves with each valu robability. This proble ice. We approach the ds. Our results are bas ast-squares model. Th	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analy ed on the large ie application is	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum s also given for general Rp/Lp —that is,						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present disci groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res	d for modeling a bina odeled using regular l ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the le- ults for large and sma	012100 ry response variable. inear regression. In o ng independent slope e consider here is the curves with each valu robability. This proble ice. We approach the ds. Our results are bas ast-squares model. Th Il samples are checke	For example, y ur work, many a s and independ construction of ue (viz. R1, R0) a em is the same problem analyt ed on the large he application is d against a 'tru	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the tically using parametric methods. A -sample properties of Maximum is also given for general Rp/Lp —that is, th source' generated using a Bootstrap						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res	d for modeling a bina odeled using regular l ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the lea ults for large and sma	012100 ry response variable. inear regression. In o ng independent slope e consider here is the curves with each valu robability. This proble ace. We approach the ds. Our results are bas ast-squares model. Th Il samples are checke	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analyti ed on the large ie application is d against a 'tru	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum s also given for general Rp/Lp —that is, th source' generated using a Bootstrap						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER	Y NOTES Center Edwards AFE s a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on re ot equal to R50. Res	d for modeling a bina odeled using regular I ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the lea ults for large and sma	012100 ry response variable. inear regression. In o ng independent slope e consider here is the curves with each valu robability. This proble ice. We approach the ds. Our results are bas ast-squares model. Th Il samples are checke	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analyti ed on the large ie application is d against a 'tru	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum s also given for general Rp/Lp —that is, th source' generated using a Bootstrap						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER Logistic, radar, R50	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res RMS D, LD50, calibratior	d for modeling a binal odeled using regular I ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the lea ults for large and sma	012100 ry response variable. inear regression. In o ng independent slope e consider here is the curves with each valu robability. This proble ice. We approach the ds. Our results are bas ast-squares model. Th Il samples are checke	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analyte ed on the large ie application is d against a 'tru	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum s also given for general Rp/Lp —that is, th source' generated using a Bootstrap						
13. SUPPLEMENTARY CA: Air Force Test of 14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER Logistic, radar, R50	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res RMS D, LD50, calibration	d for modeling a binal odeled using regular I ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the le- ults for large and sma	012100 ry response variable. inear regression. In o ing independent slope consider here is the curves with each value robability. This proble ince. We approach the ds. Our results are bass ast-squares model. The Il samples are checke	For example, y ur work, many a s and independ construction of ue (viz. R1, R0) a em is the same problem analyti ed on the large he application is d against a 'tru	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum is also given for general Rp/Lp —that is, th source' generated using a Bootstrap						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER Logistic, radar, R50 16. SECURITY CLASS	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res RMS D, LD50, calibration SIFICATION OF:	d for modeling a bina odeled using regular l ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the lea ults for large and sma	012100 ry response variable. inear regression. In o ing independent slope consider here is the curves with each value orobability. This proble ince. We approach the ds. Our results are bas ast-squares model. Th Il samples are checke t, confidence intervation	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analyti ed on the large he application is d against a 'tru al 18. NUMBER	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the tically using parametric methods. A -sample properties of Maximum is also given for general Rp/Lp —that is, th source' generated using a Bootstrap						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER Logistic, radar, R50 16. SECURITY CLASS Unclassified	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present discu- groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res RMS D, LD50, calibration GFICATION OF:	d for modeling a bina odeled using regular l ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the lea ults for large and sma	012100 ry response variable. inear regression. In o ing independent slope consider here is the curves with each value orobability. This proble ace. We approach the ds. Our results are bas ast-squares model. The Il samples are checked b, confidence intervation OF ABSTRACT	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analyti ed on the large he application is d against a 'true al 18. NUMBER OF PAGES	 = {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the tically using parametric methods. A -sample properties of Maximum s also given for general Rp/Lp —that is, th source' generated using a Bootstrap 19a. NAME OF RESPONSIBLE PERSON 412 TENG/EN (Tech Pubs) 						
 13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER Logistic, radar, R50 16. SECURITY CLASS Unclassified a. REPORT 	Y NOTES Center Edwards AFE is a non-linear metho ponses cannot be m In the present disco groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res RMS D, LD50, calibratior FIFICATION OF: b. ABSTRACT	d for modeling a bina odeled using regular l ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the le- ults for large and sma n, inverse prediction	012100 ry response variable. inear regression. In ong independent slope e consider here is the curves with each value orobability. This proble dece. We approach the ds. Our results are bass ast-squares model. The Il samples are checked as confidence intervation of ABSTRACT	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analytied on the large te application is d against a 'true al 18. NUMBER OF PAGES	 = {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum s also given for general Rp/Lp —that is, th source' generated using a Bootstrap 19a. NAME OF RESPONSIBLE PERSON 412 TENG/EN (Tech Pubs) 19b. TELEPHONE NUMBER (include area 						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER Logistic, radar, R50 16. SECURITY CLASS Unclassified a. REPORT Unclassified	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present disci groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res RMS D, LD50, calibratior FIFICATION OF: b. ABSTRACT Unclassified	d for modeling a bina odeled using regular I ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the le- ults for large and sma n, inverse prediction	012100 ry response variable. inear regression. In o ng independent slope e consider here is the curves with each valu robability. This proble ice. We approach the ds. Our results are bas ast-squares model. Th Il samples are checke ast-squares model. Th Il samples are checke the confidence interval 17. LIMITATION OF ABSTRACT None	For example, y ur work, many a s and independ construction of the (viz. R1, R0) a em is the same problem analyti ed on the large the application is d against a 'true al 18. NUMBER OF PAGES 14	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum is also given for general Rp/Lp —that is, th source' generated using a Bootstrap 19a. NAME OF RESPONSIBLE PERSON 412 TENG/EN (Tech Pubs) 19b. TELEPHONE NUMBER (include area code)						
13. SUPPLEMENTARY CA: Air Force Test (14. ABSTRACT Logistic regression is detections. Such res present themselves. comparing multiple difference in the rad represents the range LD50 ('lethal dose/e feature is the use of Likelihood estimatio range/dose values n program 15. SUBJECT TER Logistic, radar, R50 16. SECURITY CLASS Unclassified a. REPORT Unclassified	Y NOTES Center Edwards AFE a non-linear metho ponses cannot be m In the present disci groups of measures. lar 'Range 50' (R50) e at which radar ach ffective dose 50 %') 'inverse prediction' n, and improve on ro ot equal to R50. Res RMS D, LD50, calibration SIFICATION OF: b. ABSTRACT Unclassified	d for modeling a bina odeled using regular I ussion, models allowir The question that we values for two logistic ieves 50% detection p value in medical scien or calibration method esults based on the le- ults for large and sma h, inverse prediction	012100 ry response variable. inear regression. In o ng independent slope e consider here is the curves with each valu robability. This proble ice. We approach the ds. Our results are bas ast-squares model. The Il samples are checked the confidence intervation of ABSTRACT None	For example, y ur work, many a s and independ construction of ie (viz. R1, R0) a em is the same problem analyte d on the large te application is d against a 'true al 18. NUMBER OF PAGES 14	= {success, failure} for blip-scan radar applications of logistic regression ent intercepts are considered for a confidence interval about the arising from the separate curve. R50 as the problem of prediction of the cically using parametric methods. A -sample properties of Maximum is also given for general Rp/Lp —that is, th source' generated using a Bootstrap 19a. NAME OF RESPONSIBLE PERSON 412 TENG/EN (Tech Pubs) 19b. TELEPHONE NUMBER (include area code) 661-277-8615						



412th Test Wing



War-Winning Capabilities ... On Time, On Cost



Confidence Intervals for

Binary Responses-

R50 & the Logistic Model



ACAS, October 2012. Monterey, CA

Dr. Arnon Hurwitz - Edwards AFB, CA 93524

Contact: arnon.hurwitz@edwards.af.mil

Approved for public release; distribution is unlimited. 412TW-PA - 12811

Integrity - Service - Excellence

412th Test Wing

Confidence Intervals for Binary Responses ACAS 2012, Monterey, CA



U.S. AIR FORCE

Statistical Methods Group, Edwards AFB





Overview



- Blip-scan radar output returns are: detect/no-detect, or {0, 1}
- Probability of detection π increases as range-to-target decreases
- A common metric is R50 the range at which $\pi = 50\%$
- A common question is: given two flights, what is a confidence interval (C.I.) for the *difference of the two R50's ?*
- Such $R(\pi)$ differences are non-linear functions of the parameters of the estimation procedure; its own distribution is hard to derive
- A solution to find a C.I. for a difference is to use a Bootstrap procedure = a non-parametric simulation approach
- Bootstrapping works, but it has to be custom-generated for each different problem at hand. It's sometimes preferable to have a <u>parametric method</u>. We develop such a method here based on the Max. Likelihood Covariance (inverse of the Fisher Information).

We Logistic curve fit to Binary data

0

- One has the relation:
 - Output = function(Range)
- But output is binary {0, 1}, and we'd rather wish to find something like:
 - $-\pi$ = function(Range)
- Transform the problem:

 $y(R) = \log [\pi/(1 - \pi)] = \alpha + \beta R$

 Now we have a linear relation of a kind, with

 $\pi = \exp(y) / [1 + \exp(y)]$



A Logistic curve

Probability is on the vertical axis



Vertice Comparing two logistic curves

- log [π/(1 π)] == logit(π) is called the 'logit'. Log ==In
- The graph shows two such curves (two flights)
- At π = 0.5 the blue curve shows R50 at about 10 nm, the red curve at about 15 nm
- The difference is about 5: we want a 95% confidence interval around this difference



Two logistic curves

Here, P(detect) decreases with increasing R









 Let one curve be estimated with non-linear regression techniques (generalized linear modeling) to give the equation

 $logit(\pi) = \alpha_0 + \alpha_1 R$

and let the other curve be estimated as

 $logit(\pi) = \beta_0 + \beta_1 R$

• At $\pi = 0.5$, logit(π) = log(0.5/0.5) = log(1) = 0.

So for the first curve $R0 = -\hat{\alpha}_0/\hat{\alpha}_1$, and $R1 = -\hat{\beta}_0/\hat{\beta}_1$ for the second. Their estimated difference is therefore $R0 - R1 = -(\hat{\alpha}_0/\hat{\alpha}_1 + \hat{\beta}_0/\hat{\beta}_1)$

- Generalized linear modeling uses maximum likelihood estimation (MLE) techniques to estimate the coefficients of the models, and also gives us the Covariance Matrix of the α and β parameters
- Call this covariance matrix V. It is a 4x4 symmetric matrix.



Confidence Interval



• It can be shown (by MLE large-sample theory) that

 $(\widehat{R1} - \widehat{R0}) \sim Normal(R1 - R0, hVh')$

Where V is the covariance matrix, and where

$$\mathbf{h} = \left(\frac{-1}{\alpha_1}, \frac{\alpha_0}{\alpha_1^2}, \frac{1}{\beta_1}, \frac{-\beta_0}{\beta_1^2}\right)$$

This gives us the (95%) confidence interval that we desire as:

 $(\widehat{R1} - \widehat{R0}) - 1.96 \times \widehat{h}\widehat{V}\widehat{h}' < R1 - R0 < (\widehat{R1} - \widehat{R0}) + 1.96 \times \widehat{h}\widehat{V}\widehat{h}$





- So far, we've developed a CI for R50; that is, where $\pi = 0.5$
- We can get a CI for any value of π in (0, 1) by replacing the 'h' we used in the above slide with

$$\mathbf{h} = (\frac{-1}{\alpha_1}, \frac{-(y_c - \alpha_0)}{\alpha_1^2}, \frac{1}{\beta_1}, \frac{(y_c - \beta_0)}{\beta_1^2}),$$

where y_c is the estimate of the logit(π) at the new value of π .

 The above theory depends on the assumption that the two flights gave independently- estimated curves, and the curves do not cross over each other.



Bootstrap Test



We ran bootstrap simulations against our analytic technique for $\pi = p = 0.2, 0.5$, and 0.8 to see how we compared, and also compared our F.I. method against Schwenke & Milliken's

At probability	Method 1	Method 2	Method 3	Method 4
Ρ	Analytical 95% C.I. based on Fisher Information	Bootstrap 95% C.I.	Bias Corrected Bootstrap 95% C.I.	Normal approx. method of Schwenke & Milliken
0.2	[2.7, 3.7]	[2.5, 3.8]	[2.56, 3.9]	[2.4, 4.0]
0.5	[4.8, 5.9]	[4.8, 5.9]	[4.9, 5.9]	[4.5, 6.2]
0.8	[6.9, 8.0]	[6.7, 8.3]	[6.7, 8.3]	[6.7, 8.3]

Our method produces intervals very close to the Bootstrap ©

Note: These are all large-sample results





We looked at a smaller sample size (n=200), and compared our 'Analytical' method against Schwenke & Milliken's. (The data is randomly sampled from the original bootstrap data set).

p =0.5 n=200	Method 1	Method 2	Method 3
Run	Analytical 95% C.I. based on Fisher Information	Schwenke & Milliken's method 95% C.I.	Bias Corrected Bootstrap 95% C.I.
#1	[0.4, 5.5]	[-1.8, 7.8]	[4.9, 5.9]
#2	[3.8, 8.1]	[1.2, 10.8]	[4.9, 5.9]
#3	[2.1, 6.0]	[0.5, 7.6]	[4.9, 5.9]

Our method produces narrower intervals than S&M

NOTE: This conclusion is based on just 3 runs; more extensive tests are planned





- We looked at a CI on the difference between R50 points for two independent flights
- We extended the results to the difference between two Rp points, where 0
- Our 'analytic' method is based on the covariance matrix generated from the MLE procedure of generalized regression
- We compared Cl's of our method to the 'true' Cl generated by a large bootstrap sample (n = 4000 scans/flight), and also to an alternate method by Schwenke & Milliken (1991)
- We further looked at the comparative results for a 'small' sample (n = 200 scans/flight).



REFERENCES



- [1] Schwenke, J.R. and Milliken, G.A. (1991). On the Calibration Problem Extended to Nonlinear Models. Biometrics 47, 563-574
- [2] McCullagh P. and Nelder, J. A. (1989). Generalized Linear Models. Chapman & Hall
- [3] Kendall, M.G. and Stuart, A.S. (1961). *The Advanced Theory of Statistics. Vol.1*. 2nd edition. Griffin
- [4] Efron, B. and Tibshirani, R.J. (1993). An Introduction to the Bootstrap. Chapman & Hall.
- •
- [5] Hurwitz, A. (2012). Confidence Intervals for Binary Responses R50 & the Logistic Model. (AFTC: To be released)