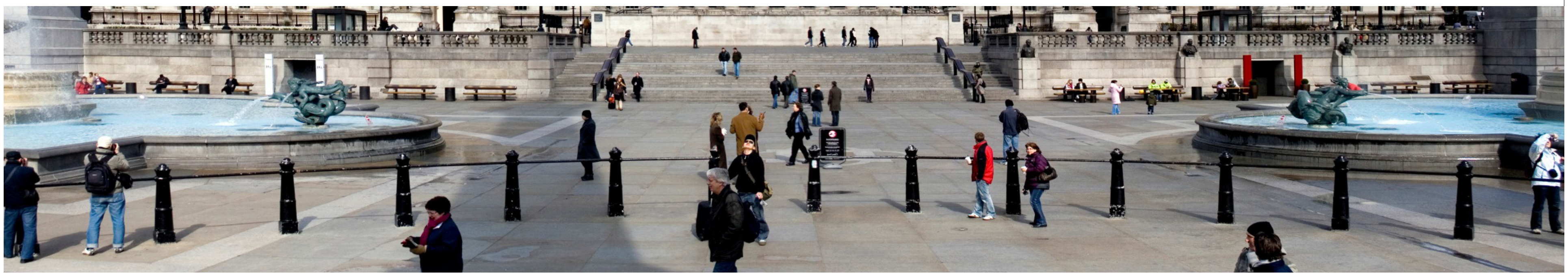
Marginal Kolmogorov-Smirnov Analysis:

Measuring Lack of Fit in Logistic Regression

Edinburgh Credit Scoring Conference
29 August 2013

Gerard Scallan
gerard.scallan@scoreplus.com

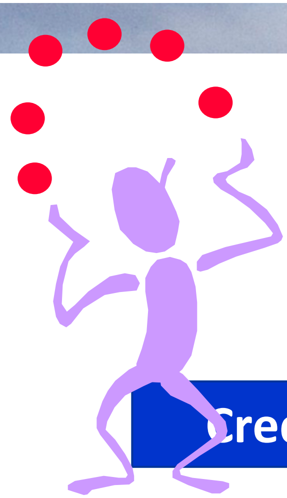
Score Plus
→ data → information → profit



Marginal Kolmogorov Smirnov Analysis
Look Ma, no grouping!

→
- Sufficient Statistic
- Marginal Information
- Marginal KS
- Open Questions

Complexity: the Enemy



Collateral

Savings

Character

Customer Lifetime Value

Affordability

Judgmental Review

Credit Bureau Derogs

+ve Credit History

Geo-demographics

Current Account Usage

Psychometrics

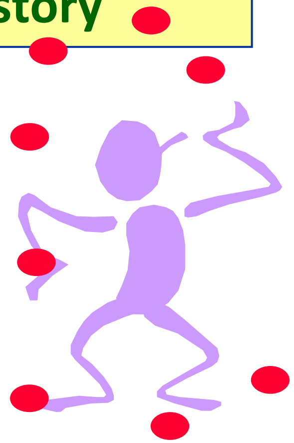
Address Validation

Price Sensitivity

Cross-Product Default

Limit Setting

Fraud Checks



Complexity → Confusion → Unmanageable → Wrong Decisions

Goal: Simplify Policies

Dimensionality Reduction → Scores

e.g. Parameter = PD

Definition:

Score = Sufficient Statistic for Aspect of Behaviour

Sufficient Statistic:

"No other statistic which can be calculated from the same sample provides any additional information as to the value of the parameter"

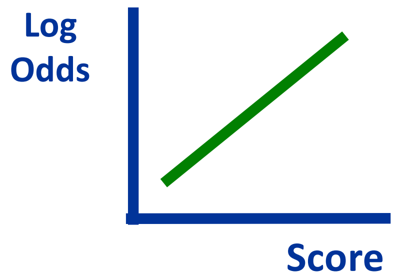
- R.A. Fisher (1922) p. 310

E.g. if Pr(Default) = 5%, then must have Pr(Default|Own) = Pr(Default|Rent) = 5%

Score Plus
→ data → information → profit

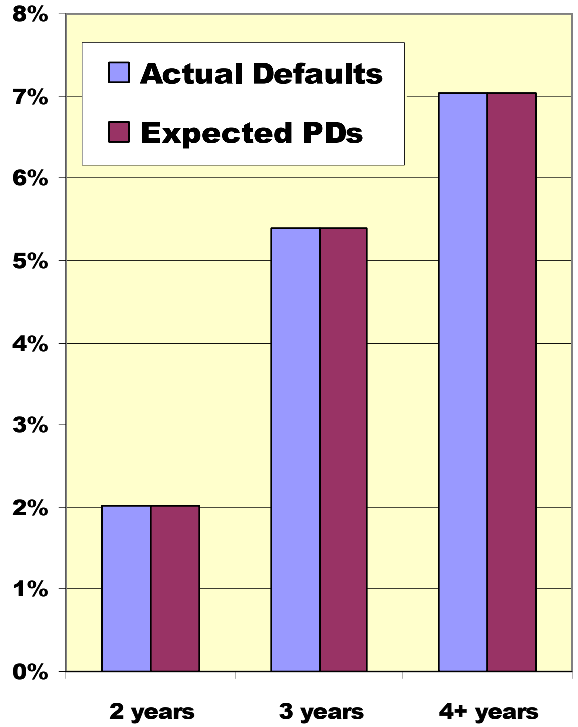
Sufficient Statistic ↔ Actual = Expected

- Sufficient Statistic → Maximum Likelihood
- \+



- → Logistic Regression
- → Actual = Expected
  - For scorecard build sample
- Proof: See Appendix

Default by Loan Term



Operational Definition of Sufficient Statistic

→ data → information → profit

Marginal Kolmogorov Smirnov Analysis
Look Ma, no grouping!

- ✓ ◆ Sufficient Statistic
- → ◆ Marginal Information
- ◆ Marginal KS
- ◆ Open Questions

Applying Actual = Expected

Development or Validation

Model Development

Given: Default Rates
Match: PDs

Model Validation

Given: PDs
Match: Default Rates

MEASURE GAP

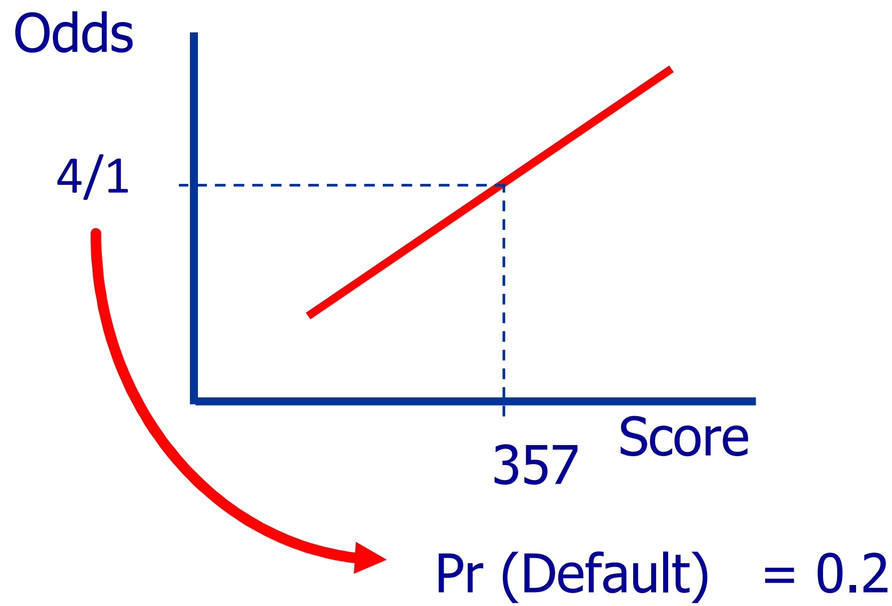
- How big?... Marginal Information
- Is it reliable?... Marginal Chi²

Modify Model
Characteristics, Attributes

Modify Model or Strategies
Score weights, new chars.
Pricing, limits, features …

Ensure Score is Sufficient Statistic

Model → Probabilities for Each Point



Measuring the Gap

Discrete Predictors

From probabilities calculated by model

Residential Status	OBSERVED				EXPECTED				
	Goods	Bads	Total	WoE	Goods	Bads	Total	WoE	Δ-score
Owner	7092	81	7173	0.639	7046.8	126.2	7173	0.189	0.450
Renter	2331	123	2454	-0.891	2376.7	77.3	2454	-0.408	-0.484
All Other	1068	23	1091	0.005	1067.5	23.5	1091	-0.017	0.022
TOTAL	10491	227	10718		10491.0	227.0	10718		
	Chi² =	42.58		D.F. = 2	p-value:	**0.00000006%**		MIV =	**0.298**

- Weight of Evidence: WoE = $\ln(G_A/B_A) - \ln(G_{total}/B_{total})$
- $Chi^2 = 2 \times \text{Observed} \times \ln(\text{Observed}/\text{Expected})$
- Degrees of Freedom: D.F. = # Attributes -1
- Δ-score = Observed WoE - Expected WoE
- Marginal Information Value: MIV = $Avg_{Goods}(\Delta\text{-score}) - Avg_{Bads}(\Delta\text{-score})$
- Cf. Scallan (2009)

Distance: Marginal Information Value

Certainty: Marginal Chi²

ScorePlus → data → information → profit

Applications of Marginal Analysis

- Selection of characteristics to enter model
- "Classing" characteristics in model
  - Now done "automatically" in logistic regression
- Model segmentation
- Policy rule assessment
- Tracking shifts in predictive patterns
- Evaluating price sensitivity

Used "everywhere" in development and validation

Score Plus
→ data → information → profit

Selecting Scorecard Characteristics

Marginal Analysis Application

		DaysXsL6m	ToB	SinceDish	AutoCr	CurDaysXs
Characteristic	***IV***	***Score1***	***Score2***	***Score3***	***Score4***	***Score5***
CurBal	0.032	0.019	0.017	0.013	0.010	0.008
CurCTO	0.185	0.121	0.086	0.089	0.007	0.006
CurDaysXs	0.616	0.125	0.113	**0.106**	**0.094**	0.021
CurDTO	0.215	0.117	0.087	0.093	0.026	0.025
CurValXs	0.515	0.121	0.110	0.093	0.090	0.007
ToB	0.692	**0.526**	0.010	0.026	0.025	0.025
MthsInact	0.012	0.005	0.001	0.004	-0.002	-0.003
MthsNoCTO	0.077	0.066	0.043	0.045	0.001	0.000
NetTO	0.074	0.028	0.007	0.010	0.002	0.000
DaysDbL3m	0.055	0.008	0.013	0.008	0.005	0.004
DaysXsL6m	**0.856**	0.000	0.008	0.011	0.015	0.012
CurMxBal	0.033	0.015	0.018	0.013	0.005	0.003
DishL1m	0.291	0.090	0.084	-0.006	-0.008	-0.010
DishL3m	0.292	0.081	0.077	0.005	0.011	0.011
SinceDish	0.810	0.397	**0.299**	0.057	0.050	**0.051**
InterCTO	0.017	0.004	-0.003	-0.004	-0.001	-0.001
InterDTO	0.003	0.001	0.000	0.000	-0.002	-0.002
AutoCr	0.209	0.143	0.108	**0.106**	0.005	0.004
ValDishL6m	0.468	0.145	0.137	-0.001	-0.001	0.003

- At each step…
- Rank predictors by (Marginal) IV
- Predictor with maximum Marginal IV enters model
- … provided Marginal Chi² can be “made” significant
- Continue until no ***significant*** Marginal IV left
- Significance Threshold:
  - MIV > .020
  - Marginal Chi² p-level < 5%

Deals with collinearity … and tells the story

Marginal Analysis as Management Tool

Model Development

Can easily analyse 1000s of predictors
… and interactions
… automate discretisation
… with reasonable computation

Coherent framework for model selection
Distance → Marginal Information
Certainty → Marginal Chi²

Ongoing Management

Can identify changes in performance
… rapidly
… and suggest corrections to model

Measure effect of policy variation
e.g. price levels, marketing mix

Evaluate policy rules

Measure consequences of overrides

Operational Implementation of Sufficient Statistic …BUT …

Issue 1A: Needs Preliminary Classing

Residential Status	OBSERVED Goods	Bads	Total	WoE	EXPECTED Goods	Bads	Total	WoE	Δ-score
Owner	7092	81	7173	0.639	7046.8	126.2	7173	0.189	**0.450**
Renter	2331	123	2454	-0.891	2376.7	77.3	2454	-0.408	**-0.484**
All Other	1068	23	1091	0.005	1067.5	23.5	1091	-0.017	**0.022**
TOTAL	10491	227	10718		10491.0	227.0	10718		
	Chi^2 =	42.58		D.F. = 2	p-value:	**0.00000006%**		MIV =	**0.298**

- Must have a preliminary classing to compute Marginal IV and Chi^2
- Scores must be calibrated on sample
  - Fit score-odds line on validation sample
  - Guarantees actual = expected for total line

Problem on continuous predictors

Issue 1B: Values Depend on Classing

Attribute	*Actual Good*	*Actual Bad*	*Expected Good*	*Expected Bad*	*Delta Score*	*Chi2 Contribution*	*MIV Contribution*
A	980	19	979	20	0.052	0.05	0.000
B	980	25	985	20	-0.228	1.18	0.006
C	980	11	971	20	0.607	4.93	0.027
D	980	29	989	20	-0.381	3.63	0.017
F	980	13	973	20	0.438	2.85	0.015
G	980	27	987	20	-0.307	2.26	0.011
H	980	17	977	20	0.166	0.48	0.002
I	980	23	983	20	-0.143	0.44	0.002
J	980	15	975	20	0.293	1.40	0.007
K	980	21	981	20	-0.050	0.05	0.000
TOTAL	9800	200	9800	200		17.27	**0.089**
	DF	*9*			*p-level*	**4.465%**	

A to C	2940	55	2935	60	0.089	0.44	0.002
D to F	2940	69	2949	60	-0.143	1.31	0.006
G to H	1960	40	1960	40	0.000	0.00	0.000
I to K	1960	36	1956	40	0.107	0.42	0.002
TOTAL	9800	200	9800	200		2.17	**0.011**
	DF	*3*			*p-level*	**53.706%**	

- Artificial data – 10 atts.
- Expecteds – no pattern
- Actuals – strong values
  - MIV = .089
  - p-level 4.4%
- But Δ-scores not sensible
- Class -> 4 attributes
  - MIV = .011
  - p-level = 53%
- Not usable!

Would you put this in a model?

Score Plus
→ data → information → profit

Issue 2: Ignores Rank Ordering

Attribute	*Actual*		*Expected*		*Delta*	*Chi2*	*MIV*
	Good	*Bad*	*Good*	*Bad*	*Score*	*Contribution*	
C	980	11	971	20	0.607	4.93	0.027
F	980	13	973	20	0.438	2.85	0.015
J	980	15	975	20	0.293	1.40	0.007
H	980	17	977	20	0.166	0.48	0.002
A	980	19	979	20	0.052	0.05	0.000
K	980	21	981	20	-0.050	0.05	0.000
I	980	23	983	20	-0.143	0.44	0.002
B	980	25	985	20	-0.228	1.18	0.006
G	980	27	987	20	-0.307	2.26	0.011
D	980	29	989	20	-0.381	3.63	0.017
TOTAL	9800	200	9800	200		17.27	**0.089**
	DF	*9*			*p-level*	**4.465%**	

- Pattern not sensible
- Re-sort to get story right
- → Pay attention to rank ordering
- … but MIV and Chi² unchanged
- … PROBLEM

Triple Test

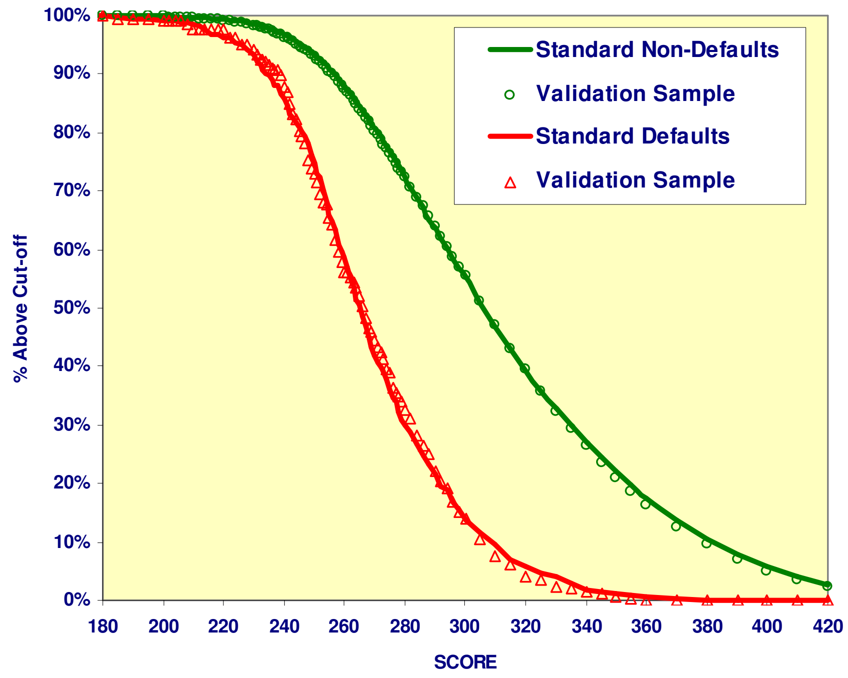
- **Distance → Marginal IV**
- **Certainty → Marginal Chi²**
- **Coherence → Tell a good story**

Marginal Kolmogorov Smirnov Analysis
Look Ma, no grouping!

- ✓ ◆ Sufficient Statistic
- ✓ ◆ Marginal Information
- → ◆ Marginal KS
- ◆ Open Questions

Find a Simpler Measure …

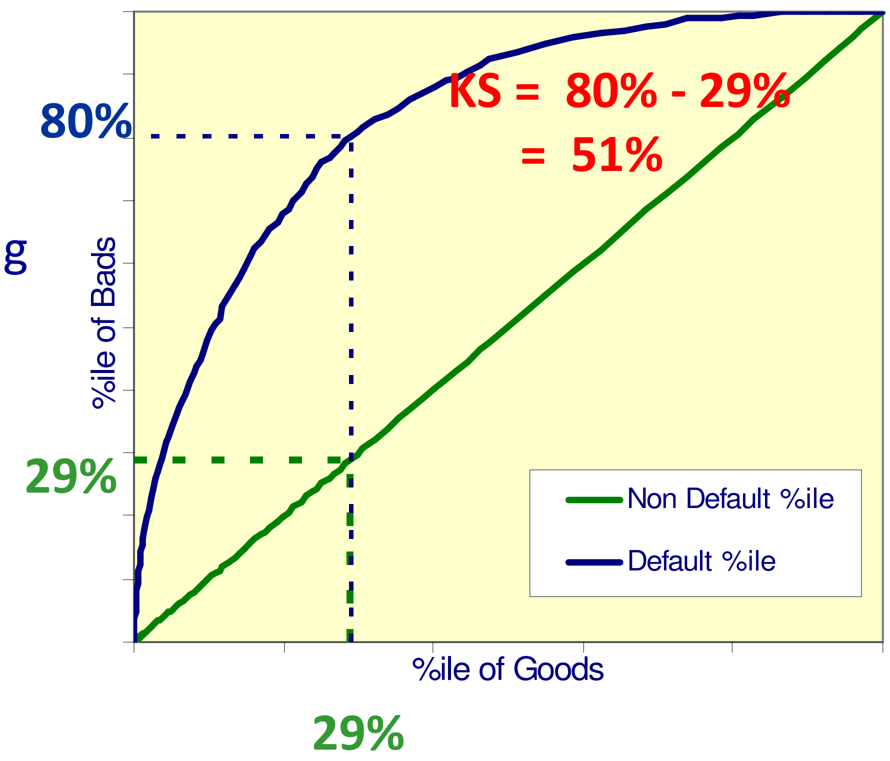
- Goal: Measure gap between "actual" and "expected"
- Continuous distribution
- → Kolmogorov Smirnov
  - Statistic
  - Test



Re-Use Familiar Concepts!

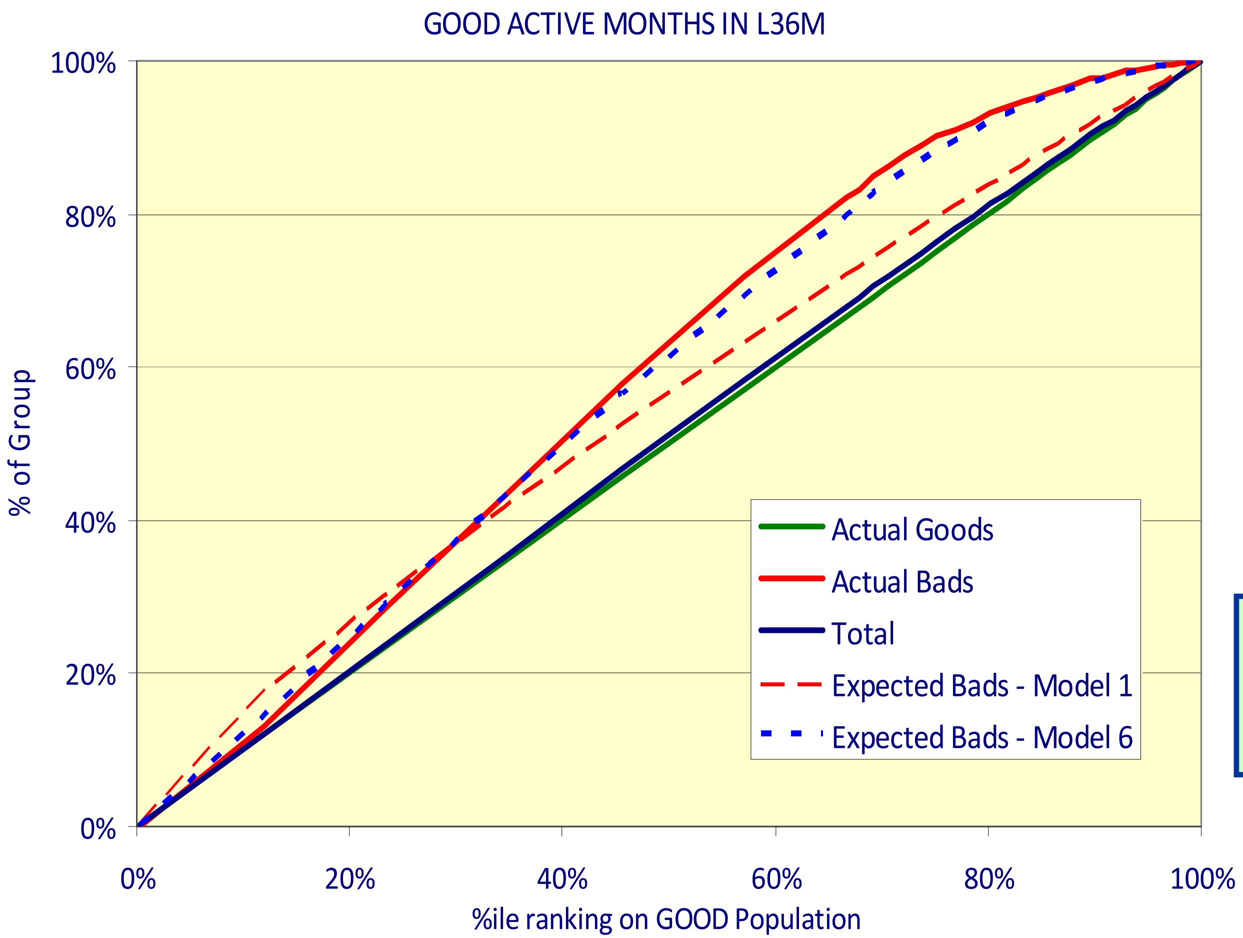
Aside: K-S and Gini Curve

- KS ranking invariant under monotone transformation of predictor
  - E.g. Score → PD – same ranking
- Convert predictor → %iles of "goods" on predictor
- Cumulative Good distribution →Diagonal!
  - By definition
- Cumulative Bads → ROC
  - a.k.a. Gini curve
- KS graph = ROC



K-S statistic is largest vertical distance between ROC and diagonal

Actual vs. Expected on Continuous Predictor



Model 0: Expected = Pop. Bad Rate

With each step "Expected" moves closer to Actual

Goal:
Expected → Actual

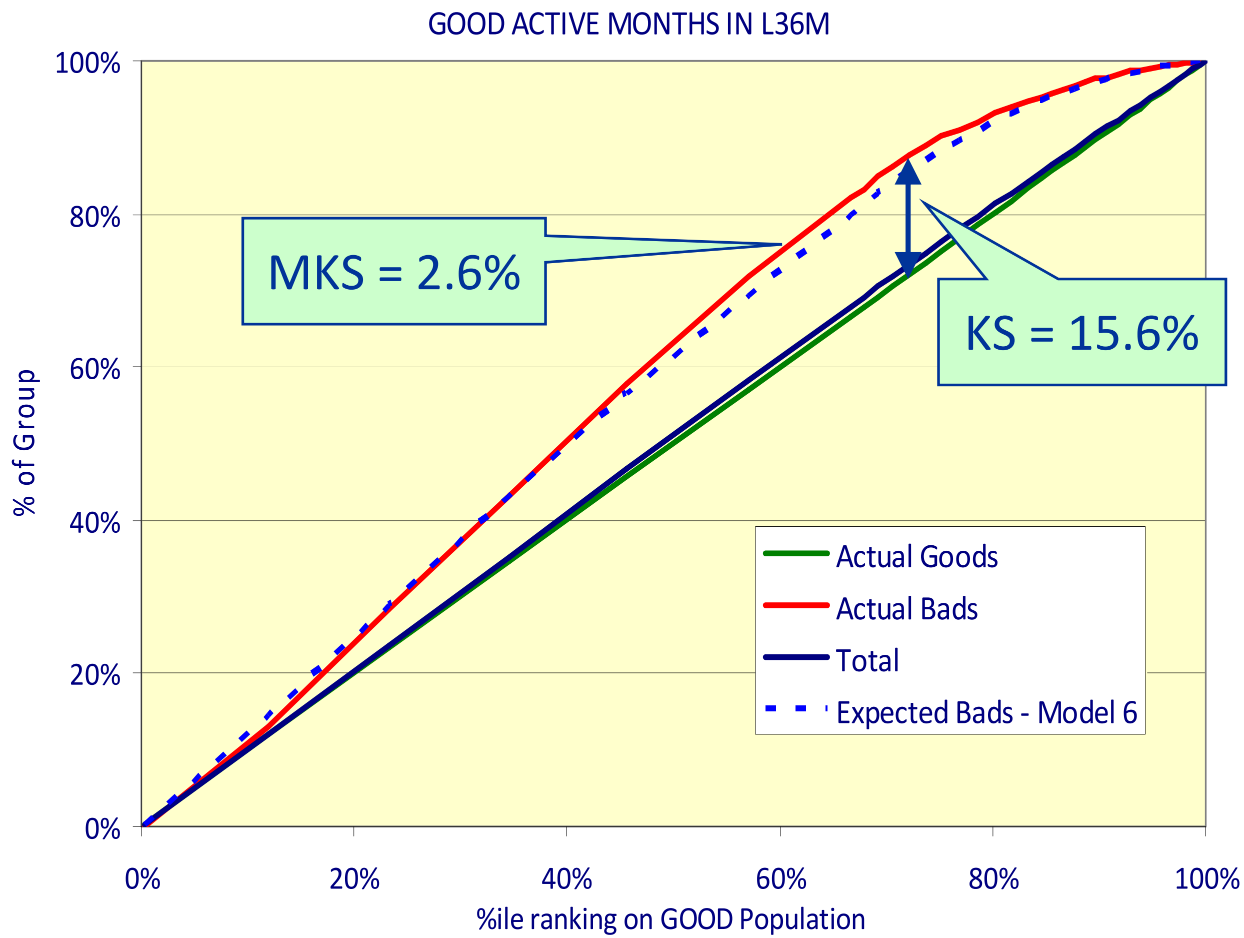
But Which KS?
Goods or Bads?

- In Numbers:
  D = (Actual – Expected Goods)
  = - (Actual – Expected Bads)
- In % terms:
  - $KS_G = \dfrac{\max |\text{Act - Exp}|}{\text{Total Goods}}$

  OR $KS_B = \dfrac{\max |\text{Act - Exp}|}{\text{Total Bads}}$
- Answer: ***Marginal KS***
  MKS = $KS_G + KS_B$
  See Appendix for details
- "Null" model: everyone has same Pr(Bad)
  = $N_B/(N_G + N_B)$
- Then: MKS = KS
- Formal Notation
  - Value of predictor i on case j = x_{ij}
  - With respect to model β
  - 1_G – indicator function
    = 1 if good, 0 if bad

$$MKS_i(\boldsymbol{\beta}) = \left(\frac{1}{N_G} + \frac{1}{N_B}\right) \max_r \left| \sum_{j:x_{ij} \le r} \left(1_G(j) - \Pr_{\boldsymbol{\beta}}(j \in G)\right) \right|$$

MKS Statistic Can Replace Marginal IV



MKS Statistic measures distance between model and reality

Marginal Kolmogorov-Smirnov Analysis
Edinburgh Credit Scoring Conference 2013

Issue 1B: Values Depend on Classing

Revisit with Marginal KS

Attribute	Actual Good	Actual Bad	Expected Good	Expected Bad	Delta Score	Chi2 Contribution	MIV Contribution	Marginal KS
A	980	19	979	20	0.052	0.05	0.000	-0.51%
B	980	25	985	20	-0.228	1.18	0.006	2.04%
C	980	11	971	20	0.607	4.93	0.027	-2.55%
D	980	29	989	20	-0.381	3.63	0.017	2.04%
F	980	13	973	20	0.438	2.85	0.015	-1.53%
G	980	27	987	20	-0.307	2.26	0.011	2.04%
H	980	17	977	20	0.166	0.48	0.002	0.51%
I	980	23	983	20	-0.143	0.44	0.002	2.04%
J	980	15	975	20	0.293	1.40	0.007	-0.51%
K	980	21	981	20	-0.050	0.05	0.000	0.00%
TOTAL	9800	200	9800	200		17.27	**0.089**	**2.55%**
	DF	*9*			*p-level*	**4.465%**		**99.96%**

Attribute	Actual Good	Actual Bad	Expected Good	Expected Bad	Delta Score	Chi2 Contribution	MIV Contribution	Marginal KS
A to C	2940	55	2935	60	0.089	0.44	0.002	-2.55%
D to F	2940	69	2949	60	-0.143	1.31	0.006	2.04%
G to H	1960	40	1960	40	0.000	0.00	0.000	2.04%
I to K	1960	36	1956	40	0.107	0.42	0.002	0.00%
TOTAL	9800	200	9800	200		2.17	**0.011**	**2.55%**
	DF	*3*			*p-level*	**53.706%**		**99.96%**

Marginal KS: independent of classing

Marginal Kolmogorov-Smirnov Analysis
Edinburgh Credit Scoring Conference 2013

→ data → information → profit

Issue 2: Ignores Rank Ordering

Revisit with Marginal KS

Attribute	*Actual*		*Expected*		*Delta*	*Chi2*	*MIV*	*Marginal*
	Good	*Bad*	*Good*	*Bad*	*Score*	*Contribution*		*KS*
A/C	980	11	971	20	0.607	4.93	0.027	-4.59%
B/E	980	13	973	20	0.438	2.85	0.015	-8.16%
C/I	980	15	975	20	0.293	1.40	0.007	-10.71%
D/H	980	17	977	20	0.166	0.48	0.002	-12.24%
A/F	980	19	979	20	0.052	0.05	0.000	-12.76%
G/K	980	21	981	20	-0.050	0.05	0.000	-12.24%
H/I	980	23	983	20	-0.143	0.44	0.002	-10.71%
B/I	980	25	985	20	-0.228	1.18	0.006	-8.16%
G/J	980	27	987	20	-0.307	2.26	0.011	-4.59%
K/D	980	29	989	20	-0.381	3.63	0.017	0.00%
TOTAL	9800	200	9800	200		17.27	**0.089**	**12.76%**
	DF	*9*			*p-level*	**4.465%**		**0.34%**

- Reminder: Marginal IV and Chi² don't change with ordering
- But Marginal KS: Unordered KS = 2.55% p-level 99.9%

  Ordered KS = 12.76% p-level 0.3%

Score Plus
→ data → information → profit

Scorecard Characteristic Selection
with Marginal Kolmogorov Smirnov

Description	MARGINAL KOLMOGOROV-SMIRNOV w.r.t. Model					
	0	**1**	**2**	**3**	**4**	**6**
Worst rating in L36m	6.8%	8.1%	6.9%	**6.0%**	3.5%	0.7%
Good Active months in L36m	15.6%	11.6%	11.9%	2.1%	2.7%	2.6%
Number of Trades at Bureau	15.5%	**15.8%**	1.0%	1.0%	1.0%	0.8%
Number of months delinquent in L36m	6.8%	8.1%	6.9%	6.0%	3.5%	2.0%
Months since last delinquent	6.8%	8.1%	6.9%	6.0%	4.2%	**0.8%**
Good - Bad Months in L48m	16.5%	12.4%	**12.8%**	2.5%	3.8%	2.8%
Good - Bad Months in L24m	15.7%	11.8%	11.8%	2.5%	3.0%	2.8%
Current Bal/Avg Bal L36m	**16.7%**	10.6%	9.0%	3.8%	4.7%	2.9%
Avg Bal L3m/Avg Bal L48m	12.9%	8.3%	6.5%	2.4%	2.5%	1.8%
Current Bal/Max Bal L6m	10.0%	5.5%	2.7%	4.8%	4.2%	0.9%
Current Bal/Avg Bal L48m (unsecured only)	12.7%	7.4%	4.6%	5.0%	**5.1%**	1.1%
Number of Revolving Accounts at Bureau L3m	15.3%	15.6%	1.3%	1.3%	1.3%	1.1%
Number of consumer credit trades at Bureau L3m	3.3%	7.0%	5.7%	2.9%	3.6%	2.9%

- Select characteristic with largest MKS to enter model
- … subject to significant MKS p-level

Marginal KS can substitute for Marginal IV

Significance Test – Kolmogorov-Smirnov

Measures Certainty

Standard KS Test

- 2-sample test
- D = max $|F_1(x) - F_2(x)|$
  - F_1 and F_2 ~ sample cumulative distn functions
- Null Hypothesis: Differences between F1 and F2 are down to random variation
  - Same underlying distribution
- Then (with sample > 50):

$$D \times \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \approx \text{ KS distribution}$$

Marginal KS Test

- Null Hypothesis: Sample comes from underlying distribution described by model
  - Discrepancies are random
- Formula for D is (a bit) different

$$D = \left(\frac{1}{N_G} + \frac{1}{N_B}\right) \max \left|\text{Actual - Exp'd}\right|$$

$$\text{so } \sqrt{\frac{N_G N_B}{N_G + N_B}} D = \sqrt{\frac{1}{N_G} + \frac{1}{N_B}} \max \left|\text{Actual - Exp'd}\right|$$

Can we believe what we see?

Model Exhausts
Predictive Content of Database

Distance

Reliability

Description	Model	0		Model	6	
	KS	*p-level*	*Info Value*	*MKS*	*p-level*	*Marg IV*
Worst rating in L36m	6.8%	0.000%	4.5%	0.7%	99.643%	0.3%
Good Active months in L36m	15.6%	0.000%	21.7%	**2.6%**	**2.181%**	**5.1%**
Number of Trades at Bureau	15.5%	0.000%	19.4%	0.8%	97.966%	0.5%
Number of months delinquent in L36m	6.8%	0.000%	4.8%	2.0%	12.593%	1.3%
Months since last delinquent	6.8%	0.000%	7.8%	0.8%	96.914%	1.7%
Good - Bad Months in L48m	16.5%	0.000%	22.9%	**2.8%**	**0.983%**	**3.9%**
Good - Bad Months in L24m	15.7%	0.000%	20.6%	**2.8%**	**1.024%**	**4.7%**
Current Bal/Avg Bal L36m	16.7%	0.000%	16.7%	**2.9%**	**0.637%**	**3.0%**
Avg Bal L3m/Avg Bal L48m	12.9%	0.000%	11.2%	1.8%	24.522%	1.9%
Current Bal/Max Bal L6m	10.0%	0.000%	5.6%	0.9%	92.980%	0.7%
Current Bal/Avg Bal L48m (unsecured only)	12.7%	0.000%	8.7%	1.1%	77.674%	0.5%
Number of Revolving Accounts at Bureau L3m	15.3%	0.000%	18.0%	1.1%	81.412%	0.7%
Number of consumer credit trades at Bureau L3m	3.3%	0.142%	7.8%	**2.9%**	**0.782%**	**2.7%**

Little information left
MKS > 2%

And mostly not reliable
p-level < 5%

Goal: Build a Sufficient Statistic

→ data → information → profit

Marginal KS: Summary

- Marginal KS gives alternative way to analyse Actual vs. Expected:

DISTANCE:	Marginal KS statistic ↔	Marginal Info Value
CERTAINTY:	MKS p-level ↔	Marginal Chi^2

- Do not need any prior classing ✓
  - But rank ordering must be reasonable …
- Coherence between Distance and Certainty measures ✓
  - Aesthetic advantage over Marginal Information/Chi^2
- Useful in development and in validation/monitoring

Useful alternative to Marginal Information

Marginal Kolmogorov Smirnov Analysis
Look Ma, no grouping!

- ✓ ◆ Sufficient Statistic
- ✓ ◆ Marginal Information
- ✓ ◆ Marginal KS
- → ◆ Open Questions

Marginal KS: Work in Progress

How to deal with "discrete attributes" in continuous predictors?
- E.g. Retired in Time on Job
- Leads to under-estimate p-levels

- Use for monitoring policy rules?

Given predictor identify model variable(s)
- Predictor > MKS point Yes/No?
- What about splines?

How is MKS affected by truncation?
- "Reject Inference" problem

Comparison of power of test
- Marginal Chi^2 vs. Marginal KS

Can MKS approximate correction to model coefficients?
- Like Δ-scores in Marginal IV

Your input is welcome!

Marginal Kolmogorov Smirnov Analysis
Look Ma, no grouping!

Thanks to:

- Banco de Credito del Perù – for the question which led to this research
  - Antonio Usquiano
  - José Carlos Sanchez
  - Elisabeth Tong
- (Anonymous) Bank which supplied test data
- Ross Gayler for persistently advocating continuous predictors

APPENDIX: THE STATISTICAL STUFF

- Actual = Expected for discrete and continuous predictors
- Weighting on Marginal KS

Actual = Expected Equations
… *equivalent to Maximum Likelihood*

Problem: estimate scorecard β from sample of Goods (G) and Bads (B)

For case i: $\Pr_\beta(i \in G) = \dfrac{e^{x_i'\beta}}{1+e^{x_i'\beta}}$ $\qquad \Pr_\beta(i \in B) = \dfrac{1}{1+e^{x_i'\beta}}$

Likelihood Function: $\quad L(\beta) = \prod_{i \in G} \dfrac{e^{x_i'\beta}}{1+e^{x_i'\beta}} \times \prod_{i \in B} \dfrac{1}{1+e^{x_i'\beta}}$

$$\ln L(\beta) = \sum_{i \in G} x_i'\beta - \sum_{i \in G \cup B} \ln(1+e^{x_i'\beta})$$

Maximise by setting partial derivatives w.r.t. each component j of β to zero:

$$\frac{\partial \ln L(\beta)}{\partial \beta_j} = \sum_{i \in G} x_{ij} - \sum_{i \in G \cup B} \frac{e^{x_i'\beta} x_{ij}}{1+e^{x_i'\beta}} = \sum_{i \in G} x_{ij} - \sum_{i \in G \cup B} x_{ij} \Pr_\beta(i \in G) = 0$$

Let $x_{ij} = 1$ if i is in category A_j, $x_{ij} = 0$ otherwise:

$$\left\| A_j \cap G \right\| \quad = \quad \sum_{i \in A} \Pr_\beta(i \in G)$$

Actual Goods = Expected Goods

Score Plus
→ data → information → profit

Actual = Expected
Continuous Variables

$$\sum_{i \in G} x_{ij} = \sum_i \frac{e^{\mathbf{x}_i'\boldsymbol{\beta}} x_{ij}}{1+e^{\mathbf{x}_i'\boldsymbol{\beta}}} \quad \text{for a continuous } x_{\cdot j}$$

$$\text{or} \quad \sum_{i \in G} x_{ij} = \sum_i \Pr_{\boldsymbol{\beta}}(i \in G) x_{ij}$$

Divide both sides by total number of "goods", $\|G\|$

$$Avg_G(x_{\cdot j}) = \frac{1}{\|G\|} \sum_i \Pr_{\boldsymbol{\beta}}(i \in G) x_{ij}$$

Because of the intercept term

$$\|G\| = \sum_i \Pr_{\boldsymbol{\beta}}(i \in G)$$

$$\text{so} \quad Avg_G(x_{\cdot j}) = \frac{\sum_i \Pr_{\boldsymbol{\beta}}(i \in G) x_{ij}}{\sum_i \Pr_{\boldsymbol{\beta}}(i \in G)}$$

$$\text{or } Avg_G(x_{\cdot j}) = Avg_{EG}(x_{\cdot j})$$

where EG is the set of "expected goods"

Average over Actuals = Average over Expecteds

How to Weight Marginal KS

Notation: $x_{ij} \sim$ value of predictor for predictor variable i, sample point j) $p_{jk} \sim$ estimated probability of "bad" (sample point j, model k)

$KS_i \sim$ Kolmogorov Smirnov statistic for predictor i $\mathrm{MKS}_{ik} \sim$ Marginal Kolmogorov Smirnov statistic for predictor i, model k

$1_{Gj}, 1_{Bj} \sim$ Indicator variables for sample point j in "Goods" or "Bads" $N_G, N_B \sim$ Total sample size of "Goods" and "Bads"

$$KS_i = \max_r \left| \sum_{j:x_{ij} \le r} \left({}^{1_{Bj}}\!/\!_{N_B} - {}^{1_{Gj}}\!/\!_{N_G} \right) \right| \qquad MKS_{ik} = \max_r \left| \sum_{j:x_{ij} \le r} \left(1_{Bj} - p_{jk} \right) \right| \times C \quad \text{for some constant } C$$

Consider MKS_{i0}, the Marginal KS w.r.t. "null" model where everyone gets same probabilities: $p_{j0} = p_0 = N_B / (N_G + N_B) \quad 1 - p_0 = N_G / (N_G + N_B)$

$$MKS_{i0} = \max_r \left| \sum_{j:x_{ij} \le r} \left(1_{Bj} - p_0 \right) \right| \times C = \max_r \left| \sum_{j:x_{ij} \le r} \left(1_{Bj} - p_0 \left(1_{Gj} + 1_{Bj} \right) \right) \right| \times C = \max_r \left| \sum_{j:x_{ij} \le r} \left(1_{Bj} (1 - p_0) - 1_{Gj} p_0 \right) \right| \times C = \max_r \left| \sum_{j:x_{ij} \le r} 1_{Bj} (1 - p_0) - \sum_{j:x_{ij} \le r} 1_{Gj} p_0 \right| \times C$$

$$= \max_r \left| \frac{N_G}{N_G + N_B} \sum_{j:x_{ij} \le r} 1_{Bj} - \frac{N_B}{N_G + N_B} \sum_{j:x_{ij} \le r} 1_{Gj} \right| \times C = \frac{N_G N_B}{N_G + N_B} \max_r \left| \sum_{j:x_{ij} \le r} {}^{1_{Bj}}\!/\!_{N_B} - \sum_{j:x_{ij} \le r} {}^{1_{Gj}}\!/\!_{N_G} \right| \times C$$

This reduces to KS_i iff: $C = \dfrac{N_G + N_B}{N_G N_B} = \dfrac{1}{N_G} + \dfrac{1}{N_B}$

Marginal KS Statistic: $$MKS_{ik} = \left(\frac{1}{N_G} + \frac{1}{N_B} \right) \max_r \left| \sum_{j:x_{ij} \le r} \left(1_{Bj} - p_{kj} \right) \right|$$

References

- R.A. FISHER (1922) “On the Mathematical Foundations of Theoretical Statistics” (Philosophical Transactions of the Royal Society, A, 222: pp. 309-368)
- Gerard SCALLAN (2009) “Marginal Chi² Analysis: Beyond Goodness of Fit for Logistic Regression Models” (http://www.scoreplus.com/ref/001.pdf)
- W. J KRZANOWSKI and D.J. HAND (2011) “Testing the difference between two Kolmogorov-Smirnov values in the context of Receiver Operating Characteristic curves” (Journal of Applied Statistics, 38, pp. 437-450)