	Bias Discussion 2.0 – Moving from Theory to Practice	
	CAS RPM	
	San Diego, California 15 March 2023	
	Christopher Cooksey, Guidewire Software Senior Director of Advanced Analytics	
1		
	Agenda	
•	Options for checking and intervening	
	Correlations are tricky	
	3. BOP example	
2		
	Where to Intervene?	
	Historical Data Analytical Implementation Business Processes	
	Point of evaluation  Participatel, But from an analytics	
	Model     Univariates     Normalization     Normalization  Model Output  Output	
	Point of evaluation and intervention intervention  - Fairness metrics  Ensure the model is used	
	Consider productors used  Minimize chance of provies  Control variables  Control variables	
3		

Morldina	fram	V-	مامما
Working	110111	KIIOW	ieuge

At the beginning of a modeling project, what is the univariate pattern across protected classes, given the target objective?







4

### Effects of correlation

360 cells, each homogeneous with respect to the target.

To model this, we identify attributes that usefully discriminate.

1.801	1.732	1.601	1.539	1.480	1.369	1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000	
1.732	1.665	1.539	1.480	1.423	1.316	1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000	
1.665	1.601	1.480	1.423	1.360	1.265	1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1
1.601	1.539	1.423	1.369	1.316	1.217	1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	0.980	1
1.539	1.480	1.369	1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	0.980	
1.480	1.423	1.316	1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000	1.000	0.980	0.960	
1.423	1.369	1.265	1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	1
1.369	1.316	1.217	1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	1
1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	1
1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	1
1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	1
1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	
1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	1
1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	1
1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	
1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	1
1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	1
1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	
1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	
1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	0.785	

5

# Effects of correlation

Attribute 1 identifies three equalsized groups.

Averaging the experience gives a useful spread in the target.

1.948	1.873	1.732	1.665	1.601	1.480	1.423	1.369	1.265	1.217	1.170	1.082	1.040	1.000	1.000
1.873	1.801	1.665	1.601	1.539	1.423	1.369	1.316	1.217	1.170	1.125	1.040	1.000	1.000	1.000
1.801	1.732	1.601	1.539	1.480	1.369	1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000
1.732	1.665	1.539	1.480	1.423	1.316	1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000
1.665	1.601	1.480	1.423	1.369	1.265	1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000
1.601	1.539	1.423	1.369	1.316	1.217	1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	0.980
1.539	1.480	1.369	1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	0.980
1.480	1.423	1.316	1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000	1.000	0.980	0.960
1.423	1			1.170	1.082				1.000	1.000				0.960
1.369	1			1.125	1.040				1.000	1.000				0.960
1.316	1	1.15	9	1.082	1.000		0.973		1.000	1.000		0.868		0.941
1.265	1			1.040	1.000				1.000	1.000				0.922
1.217	1.00	4.004	4.000	1.000	1.000	4.000	2000	4.000	1.000	0.980	0.000	0.044	0.366	0.904
1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886
1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868
1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868
1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851
1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834
1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817
1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801
1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801
1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	0.785
1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	0.785	0.769

ects	Ωf	cor	rola	stion

1.217 1.170 1.125 1.062 1.040 1.000 Attribute 2 1.082 1.040 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.960 0.961 0.964 0.964 0.886 0.888 0.881 0.834 1.040 1.000 1.000 1.000 1.000 1.000 1.000 0.980 0.980 0.980 0.980 0.980 0.981 0.886 0.886 0.886 0.886 0.881 0.834 0.817 0.817 1.000 1.000 1.000 1.000 1.000 0.980 0.960 0.960 0.961 0.962 0.963 0.863 0.863 0.863 0.863 0.863 0.863 0.863 identifies three equalsized groups, as well, but 1.000 1.000 1.000 1.000 1.000 0.980 0.960 1.000 1.000 1.000 1.000 0.980 0.960 0.941 0.922 0.922 0.904 0.886 0.868 0.861 0.861 1.316 1.265 1.217 1.170 1.125 1.082 1.040 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.265 1.217 1.170 1.125 1.082 1.040 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1217 1170 1125 1082 1040 1000 1000 1000 1000 1000 0.980 0.980 splits the data differently. Averaging the 0.960 0.941 0.941 0.922 0.904 0.886 0.868 0.868 1.082 1.040 1.000 1.000 1.000 1.000 experience gives a useful spread in the target.

7

### Effects of correlation

What if there is a protected class to keep track of?

And there is a difference in the data between groups.



8

#### Effects of correlation

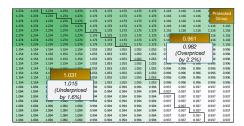
The uneven distribution changes the averages.

Which group is over or underpriced has flipped.

2.026	1.948	1.801	1.732	1.005	1.539	1.480	1.423	1.316	1.265	1.217	1.125	1.082		
1.948	1.873	1.732	1.665	1.601	1.480	1.423	1.369	1.265	1.217	1 170	1.082	1,040		otected
1.873	1.801	1.665	1.601	1.539	1.423	1.369	1.316	1.217	1.17		164	.000	- 1	Group
1.801	1.732	1.601	1.539	1.480	1.369	1.316	1.265	1.170	1,73				1.000	1.000
1.732	1.665	1.539	1.480	1.423	1.316	1.265	1.217	1.125	1.	0.9	909	10	1.000	1.000
1.665	1.601	1.480	1.423	1.369	1.265	1.217	1.170	1.082	1.	(Unde	mrica	d o	1.000	1.000
1.601	1.539	1.423	1.369	1.316	1.217	1.170	1.125	1.040	1.			u 10	1.000	0.980
1.539	1.480	1.369	1.316	1.265	1.170	1.125	1.082	1.000	1	by 5	.4%)	10	1.000	0.980
1.480	1,423	1.316	1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000	1.000	0.980	0.960
1.423	1			1.170	1.082				1.000	1.000				0.960
1.369	1	1.15		1.125	1.040		0.973		1.000	1.000		0.86		0.960
1.316	1	1.10	9	1.082	1.000		0.973		1.000	1.000		U.00	9	0.941
1.265	1			1.040	1.000				1.000	1.000				0.922
1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904
1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886
1.125	1.082	1.000				1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868
1.032	1.040	1.000		1.031	b	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868
1.040	1.000	1.000			_	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851
1.000	1.000	1.000		1.073		0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834
1.000	1.000	1.000	(Ov	erprice	ed	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817
1.000	1.000	1.000				0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801
1.000	1.000	1.000	Dy	4.0%	/	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801
1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	0.785
1.000	1.000	0.090	0.060	0.041	0.022	0.004	0.000	0.000	0.051	0.004	0.917	0.901	0.796	0.760

_		
	Effects of correlation  Note the double-whammy in the high freq group: the protected group is over-represented AND systematically lower risk.	
	Attribute 2 is 2006 1948 1893 1792 1666 1539 1480 1493 1366 1227 1225 1082 1948 1893 1792 1666 1600 1480 1480 1385 1287 1170 1082 1040 Group	
	correlated 2372 865 1539 1440 1451 1355 1516 1517 1518 1518 1518 1518 1518 1518 1518	
	protected 1480 1493 1316 1285 1297 1105 1082 1040 1000 100 100 1.00 1.00 1.00 1.00 1	
	But only 125 106 100 100 100 100 100 100 000 000 000	
	Serves to 100 100 100 107 107 100 100 100 100 100	
	1000 1000 1000 1000 6000 6000 6000 6000	
10		
	Effects of correlation – Lessons?	
	1. Averages do funny things as	
	1. Averages do funny things as the distribution changes. 2. Excluding information and being blind to the relevant attributes does not help. 3. It is not obvious ahead of time the group impact for a given	
	2. Excluding information and being blind to the relevant attributes does not help.  3. It is not obvious ahead of time the group impact for a given	
	2. Excluding information and being blind to the relevant attributes does not help.  3. It is not obvious ahead of time the group impact for a given	
	Simple models distort more. It is instructive to see what happens if we look at	
	each attribute as a single-variable model, and then do a simple ensemble.	
11		
	Effects of correlation – Lessons?	
	1204 1274 1274 1274 1274 1274 1271 1272 1273 1273 1273 1275 1275 1276 1276 1276 1276 1276 1276 1276 1276	
	1296   1296   1296   1296   1296   1296   1297   1297   1297   1297   1298   1396	
	1.50* 2.50* 1.51* 1.54 1.55 1.50 1.003 1.03 1.03 1.03 1.03 1.03 0.05 0.05 0.06 0.06 0.06 0.06 0.06 1.00 1.50* 1.50* 1.50* 1.50* 1.50* 1.50* 1.50* 1.50* 1.50* 1.50* 1.50* 1.00* 0.06 0.06 0.06 0.06 0.06 0.06 0.06	
	1284 5.548 1.548 1.548 1.554 1.554 1.555 1.550 1.503 1.504 1	
	1398 1398 1398 1398 1398 1398 1398 1398	

Effects of correlation - Lessons?

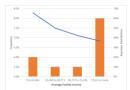


13

## BOP Example (Property coverage; Frequency target)

Using the Zip-level field "Average Family Income" as our measure of bias. Lower incomes areas show higher property frequencies.

- Naïve GLM 10 predictors; no use of an income field.
- 2. Controlled GLM added income field to control for its effect.
- Normalized GLM first normalized the data to remove the variation in frequency across income; then fit the 10-predictor GLM.



Note: these results are specific to one book of business for one target and should not be seen as generalizable. These GLMs were created quickly as illustrations, not as the best possible models of the data.

14

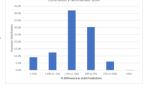
### BOP Example (Property coverage; Frequency target)

Interestingly, the Controlled and Normalized GLMs give surprisingly different results.

In the graph we compare model predictions record by record.

Remember, normalizing data gives 100% credibility to the univariate pattern.

Avg Family Income	Exposure Distribution	Univariate Freq Rel	Controlled GLM Freq Rel
0 to 65,464	20%	1.212	1.130
65,465 to 69,771	10%	1.098	1.059
69,772 to 73,236	10%	1.042	0.993
73,237 or more	60%	1.000	1.000



<b>BOP Example</b>	(Property	coverage;	Frequency	v target
--------------------	-----------	-----------	-----------	----------

Does the Controlled GLM indicate a problem with the Naïve GLM? One approach is to see if other fields lost predictive power.

Predictors 9 and 7 had the largest drop, but nothing eye-popping. When included, Income had a low influence on the model.

There's always more than just the numbers. Do you get more curious when you note Pred7 is Protection Class?

	Naïve GLM	Controlled GLM	
	Mean Absolute	Mean Absolute	
Features	SHAP Values	SHAP Values	Shift
Predictor9	0.0393	0.0396	0.0002
Predictor2	0.0162	0.0157	-0.0005
Predictor1	0.0096	0.0095	-0.0001
Predictor4	0.0070	0.0070	0.0000
Predictor5	0.0060	0.0060	0.0000
Predictor10	0.0051	0.0051	0.0000
Predictor7	0.0045	0.0042	-0.0003
Predictor3	0.0041	0.0042	0.0001
Predictor8	0.0039	0.0043	0.0004
Ave Family Income	na	0.0031	na
Predictor6	0.0028	0.0028	0.0000

16

## BOP Example (Property coverage; Frequency target)

Does the Controlled GLM indicate a problem with the Naïve GLM? Another approach is look at the changes in the fitted coefficients.

Again, nothing dramatic pops, but let's look more closely at the two predictors already identified – a Building/Contents Indicator and Protection Class.

Building and contents	40.9%	1.492	1.573	1.549	1.659	-1.5%
Building only	13.0%	0.958	1.043	1.037	1.135	-0.6%
Contents only	44.9%	1.000	1.000	1.000	1.000	0.0%
No building or contents	1.2%	0.468	0.765	0.758	1.250	-0.9%
Protection Class	Expo Dist	GLM1-low	GLM1 Rel	GLM2 Rel	GLM1-high	% Decrease
Protection Class unknown	Expo Dist 1.1%	GLM1-low 0.553	0.948	0.935	GLM1-high 1.626	% Decrease -1.4%
unknown	1.1%	0.553	0.948	0.935	1.626	-1.4%

BuildingContents Indic Expo Dist GLM1-low GLM1 Rel GLM2 Rel GLM1-high % Decrease

17

#### BOP Example (Property coverage; Frequency target)

Are the differences here large? No.

Are the differences statistically significant? Not even close.

Should you leave here today thinking I've shown some pattern that holds in general for other books of business? Please, no.

If you do this analysis and see something like this, is it worth thinking more about? I think so.

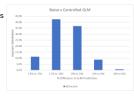
F	rotection Class	Expo Dist	GLM1-low	GLM1 Rel	GLM2 Rel	GLM1-high	% Decrea:
	unknown	1.1%	0.553	0.948	0.935	1.626	-1.4%
	0-2 (best)	22.6%	0.871	0.928	0.921	0.988	-0.7%
	3	27.1%	1.000	1.000	1.000	1.000	0.0%
	4	18.6%	0.971	1.031	1.032	1.094	0.1%
	5	10.8%	1.080	1.156	1.138	1.237	-1.5%
	6+ (worst)	19.9%	1.052	1.120	1.096	1.191	-2.1%

### BOP Example (Property coverage; Frequency target)

Does the Controlled GLM indicate a problem with the Naïve GLM? How about the model output? Do the predictions show anything?

The negative (left) end of the chart is when the Naïve GLM prediction is lower than the Controlled GLM prediction.

Overall, the predictions are quite similar. This is good.



19

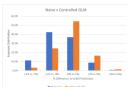
## BOP Example (Property coverage; Frequency target)

Does the Controlled GLM indicate a problem with the Naïve GLM? How about the model output? Do the predictions show anything?

With model output, you can filter down to the group we are interested in checking.

This shows that lower income areas see a higher frequency prediction at a disproportionate rate when a naïve approach is used.

Big? No. There? Yes. (in this book, this example...)



20

# BOP Example (Property coverage; Frequency target)

Thoughts on working through this example:

Checking the correlation with Income wouldn't have been sufficient.

Predictor	Correl	SHAP	SHAP Change	Change in Rels
Predictor8	14.8%	0.0039 (second lowest)	+0.0004 (biggest increase)	One level out of 10 saw a 2% move in relativity. Nothing systematic.
Protection Class	14.7%	0.0045	-0.0003 (second biggest decrease)	And yes, an interesting, though small, movement in the relativites.
State	11.0%	0.0028 (lowest)	No change	One state shifted 2%.
Building/ Contents	Not even	0.0162 (second	-0.0005 (biggest	One level dropped 2%.

- Checking multiple measures correl, SHAP, model output distributions gives much more context.
- $\bullet\ \$  In the end, this requires thought. Doubtful about using automatic measures.

Questions?	
Christopher Cooksey	
Senior Director of Advanced Analytics ccooksey@quidewire.com	
Guidewire Software	