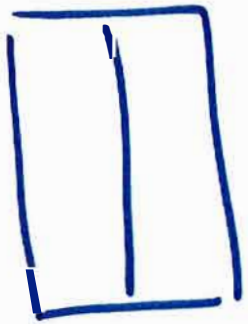
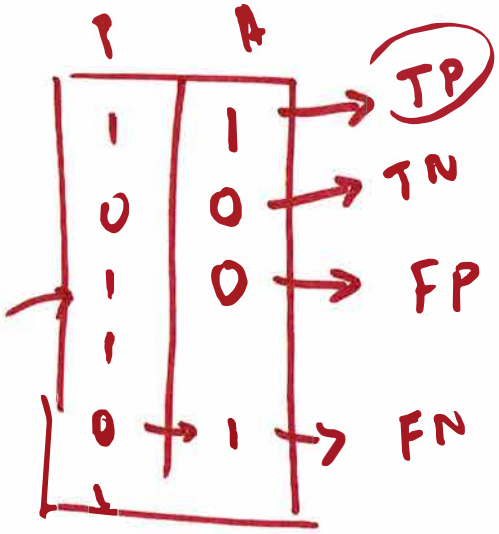
Performance Measures

greatlearning
Learning for Life

- Confusion Matrix
- ROC Curves, Gini Coefficient
- Gain and Lift Chart
- Kolomogorov-Smirnov (K-S) chart
- Concordance-Discordance ratio
- Root Mean Square Error, Mean Absolute Error

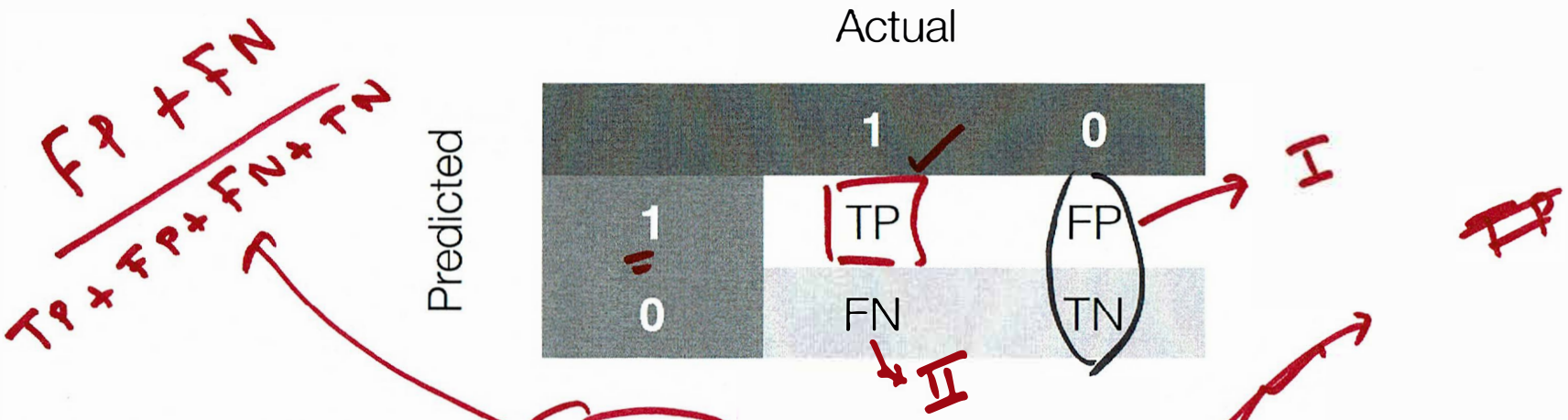




Confusion Matrix

greatlearning
Learning for Life

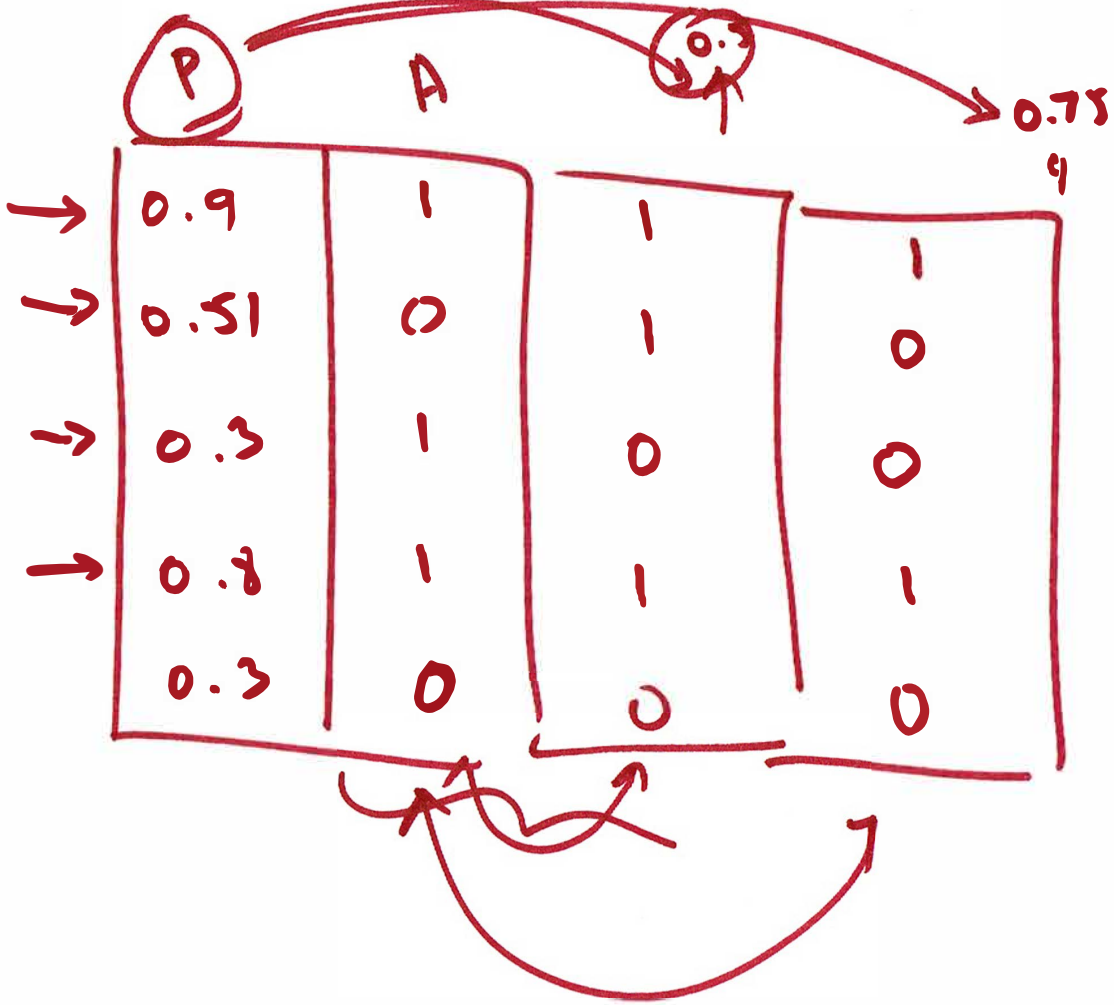
- For classification problem with a class output, the confusion matrix gives the counts of correct and erroneous predictions:

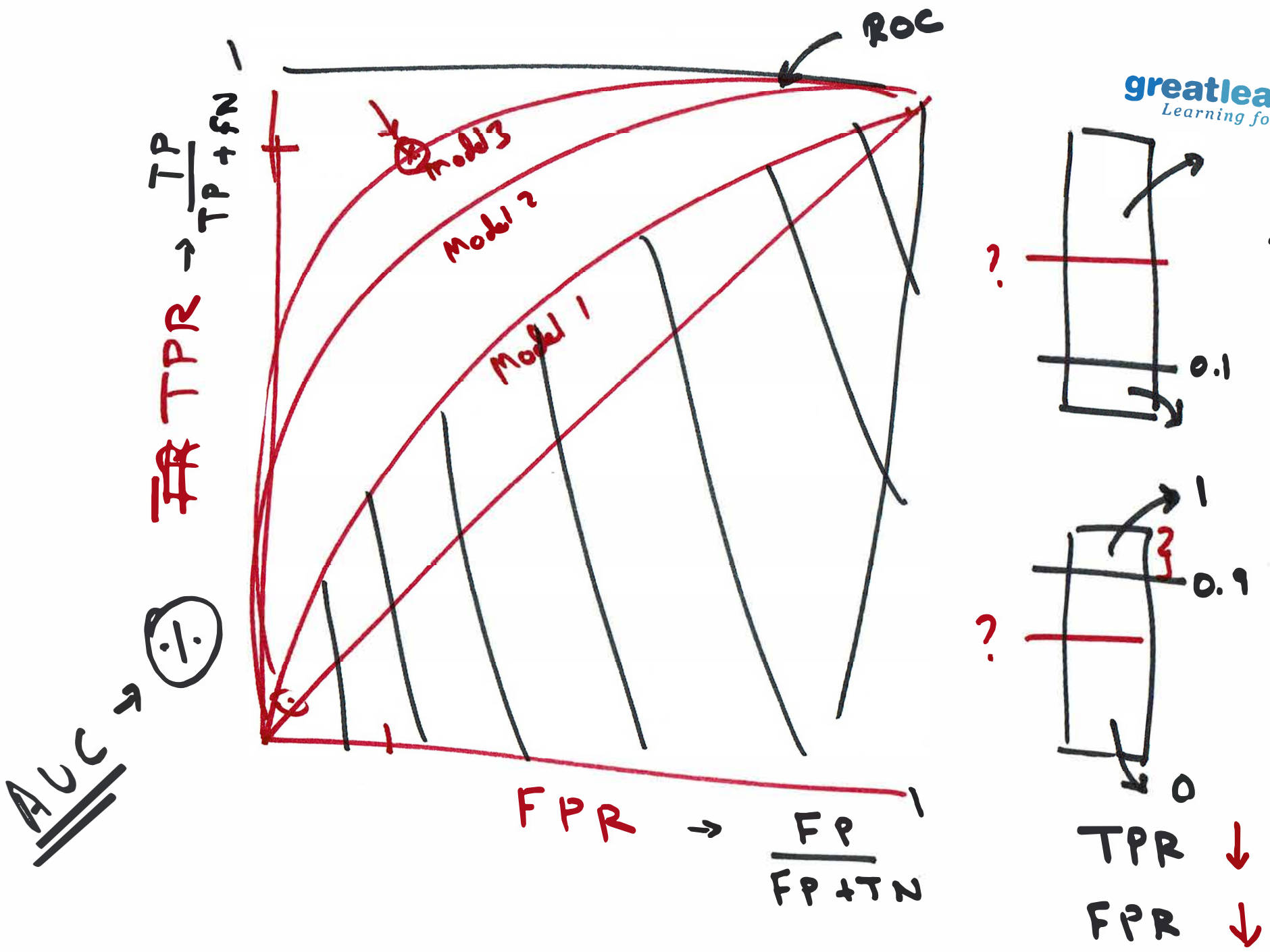


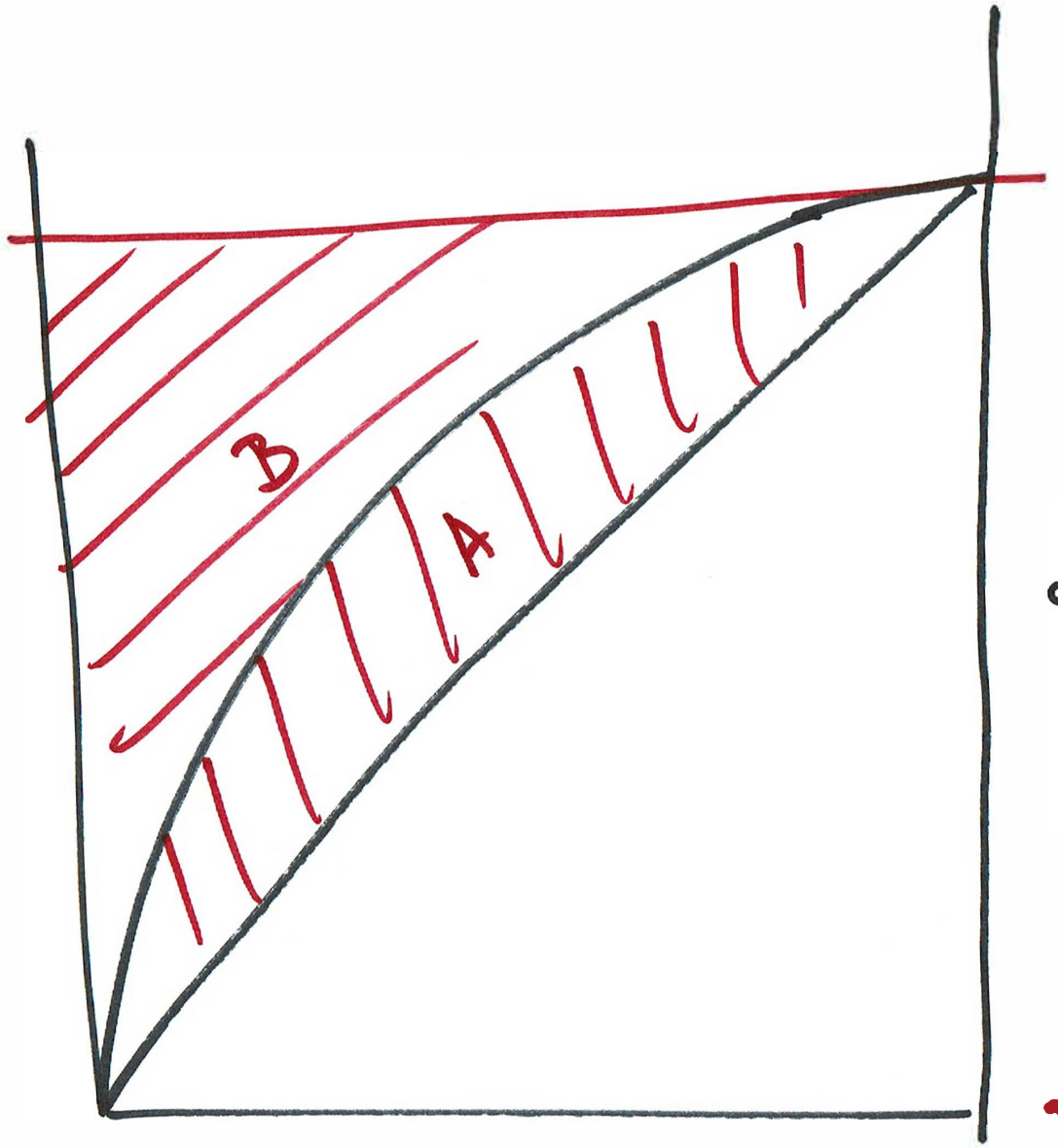
- Classification Error Rate: sum of Type 1 (FP) and Type 2 (FN) Errors (in percentage). Accuracy is 1-(error rate)

- Sensitivity (also called Recall or True Positive Rate): proportion of Total Positives that were correctly identified

- Specificity (also called True Negative Rate): proportion of Total Negatives that were correctly identified







~~AC~~

AUC

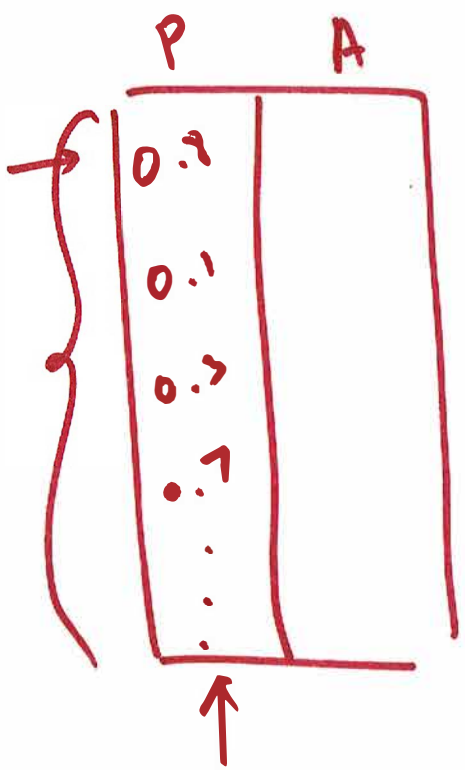
gini coeff = $\frac{A}{A+B}$

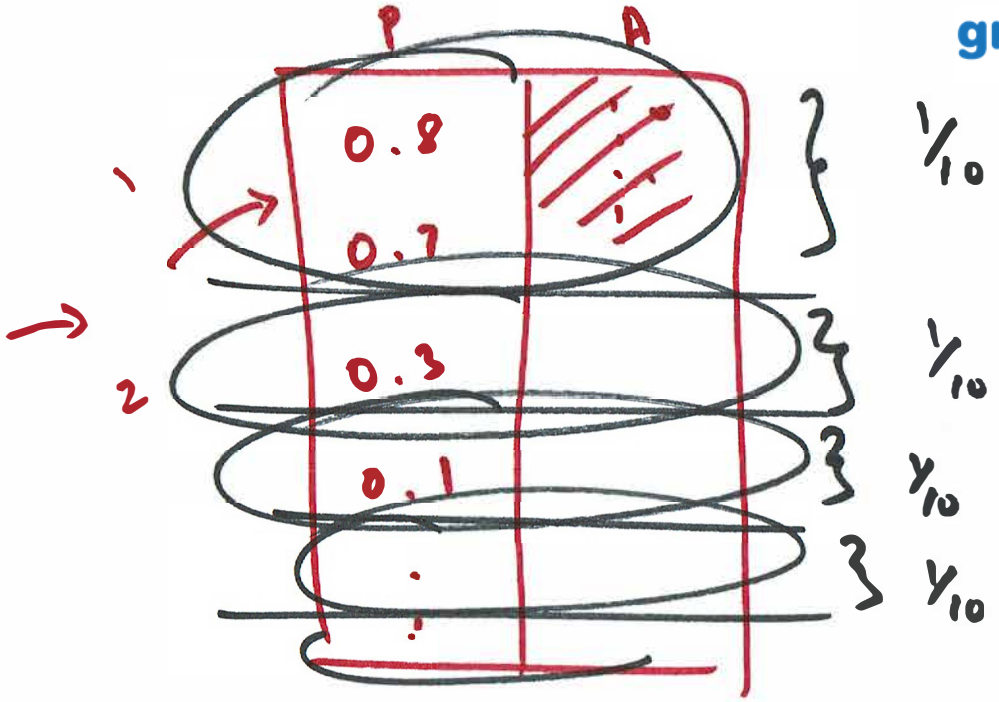
$A = AUC - 0.5$

$A + B = 0.5$

gini coeff = $\frac{AUC - 0.5}{0.5}$

$$Gini = 2AUC - 1$$

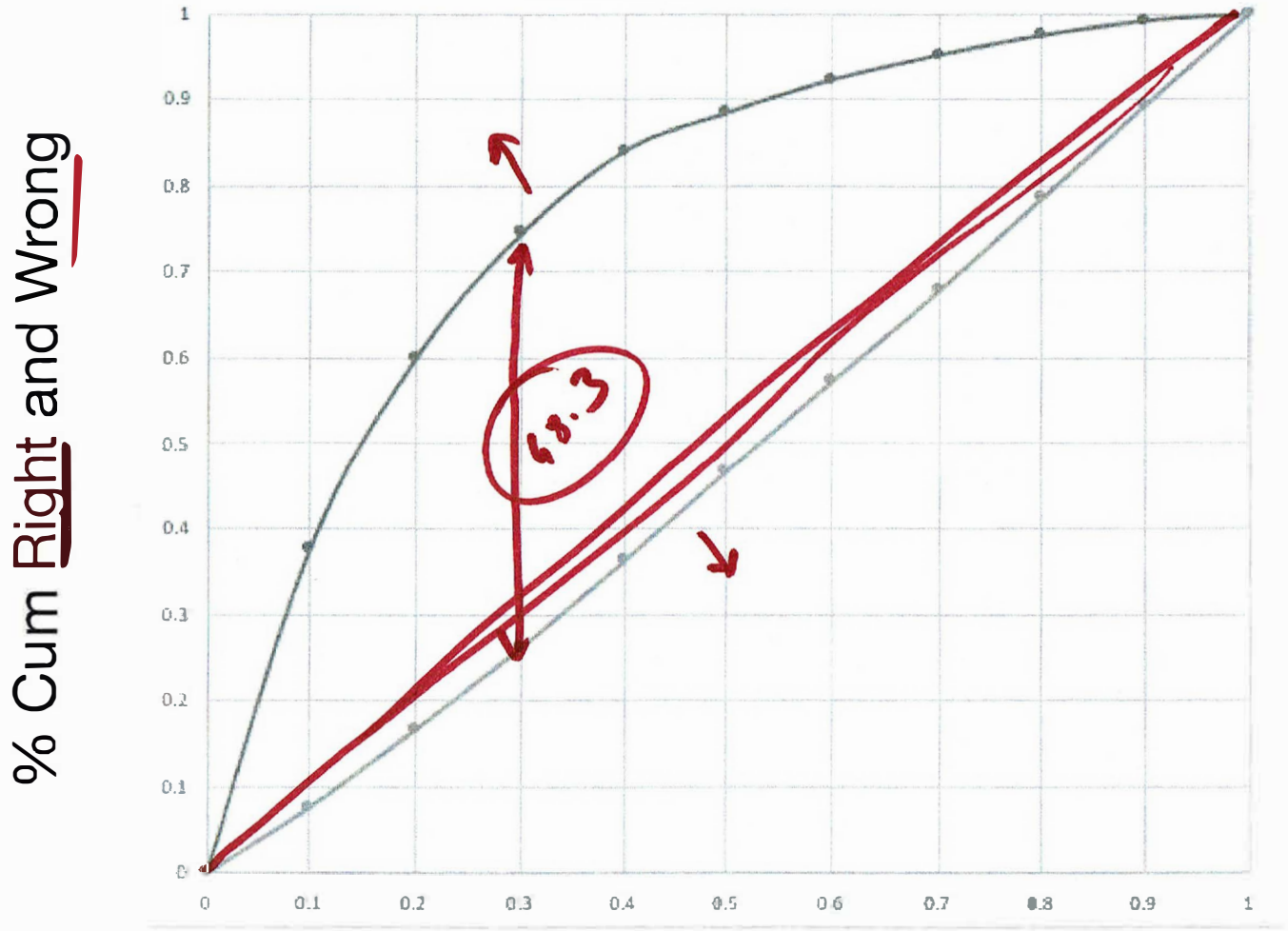




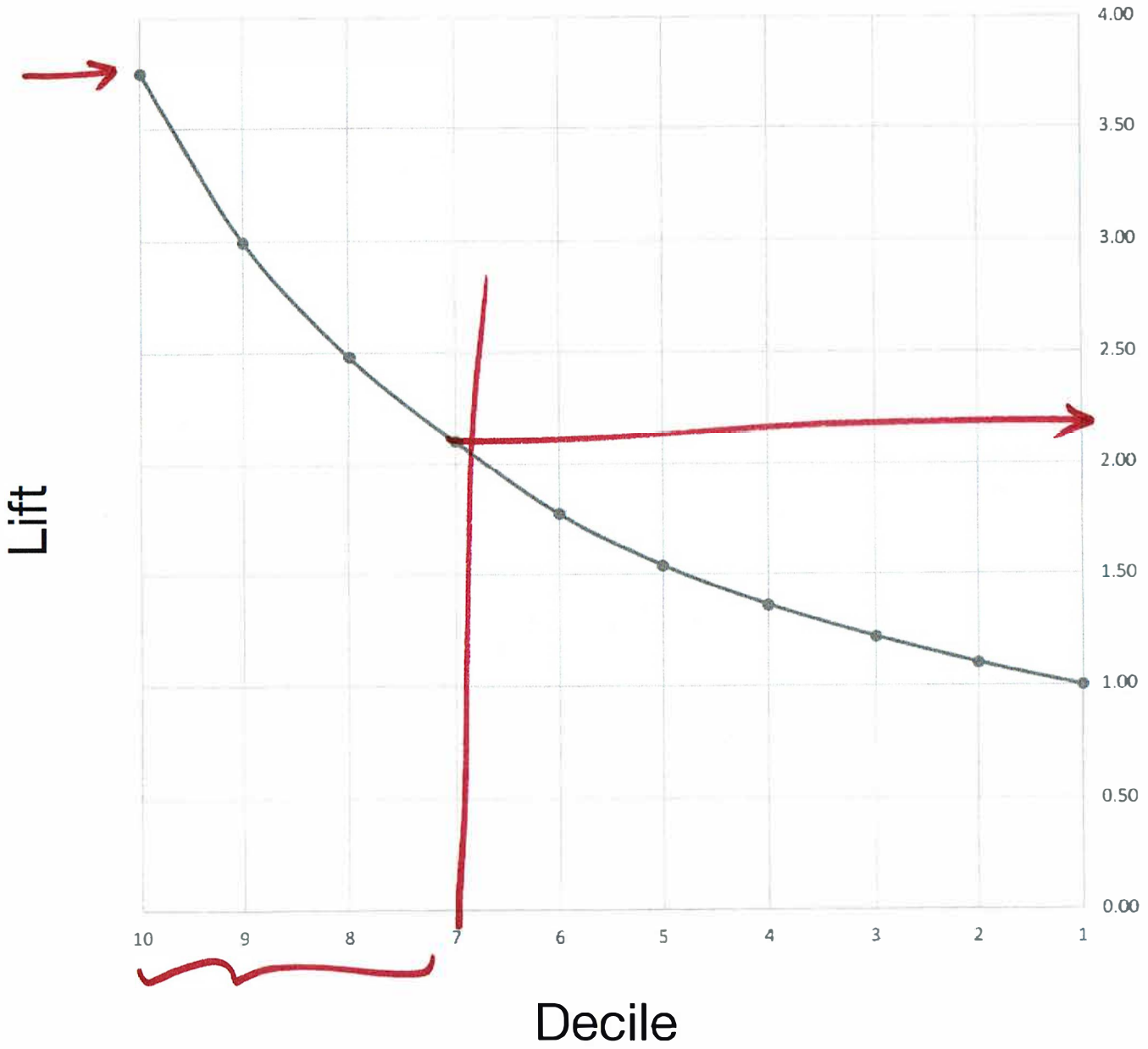
Rank Ordered Table Example

Decile	Base Cnt	#Right	# Wrong	%Right	Cum. Resp.	Cum. Non-	%Cum . Base	%Cum . Resp.	%Cum . Non-	KS	Lift
A	B	C	D=B-C	E=C/B	F = CumSum(C)	G=CumSum(D)	H = CumSum(B)/ Total	I = F/Total	J=G/Total	I-J	I/H
10	1000	295	705	29.50%	295	705	10%	37.48%	7.65%	29.83%	3.75
9	1000	176	824	17.60%	471	1529	20%	59.85%	16.60%	43.25%	2.99
8	1000	115	885	11.50%	586	2414	30%	74.46%	26.20%	48.26%	2.48
7	1000	75	925	7.50%	661	3339	40%	83.99%	36.24%	47.75%	2.10
6	1000	35	965	3.50%	696	4304	50%	88.44%	46.72%	41.72%	1.77
5	1000	30	970	3.00%	726	5274	60%	92.25%	57.25%	35.00%	1.54
4	1000	23	977	2.30%	749	6251	70%	95.17%	67.85%	27.32%	1.36
3	1000	18	982	1.80%	767	7233	80%	97.46%	78.51%	18.95%	1.22
2	1000	13	987	1.30%	780	8220	90%	99.11%	89.22%	9.89%	1.10
1	1000	7	993	0.70%	787	9213	100%	100.00%	100.00%	0.00%	1.00
Total	10000	787	9213	7.87%	787	9213					

K-S Chart



Lift Chart



Example

Name	Right?	Prob
A	0	0.056
B	0	0.134
C	0	0.156
D	1	0.512
E	0	0.235
F	0	0.25
G	1	0.25
H	1	0.2
I	0	0.135
J	0	0.089

A P

D, A → 0.512, 0.056 → ✓ C

D, B → 0.512, 0.134 → ✓ C

D, C

D, E

D, F

D, I

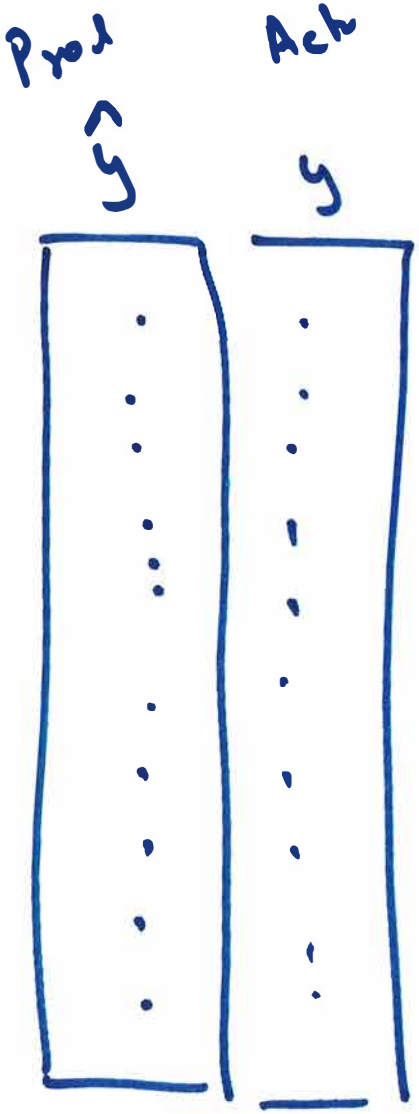
D, J

G, A

G, F → 0.25 0.25 → D T

H, E → 0.2, 0.235 → X D

Concordance Ratio = $\frac{18}{21}$



MAE

$$\frac{1}{N}\sum |y_i - \hat{y}_i|$$

RMSE

$$\sqrt{\frac{1}{N}\sum (y_i - \hat{y}_i)^2}$$