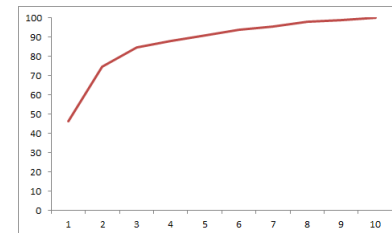


> 모형 평가 방법 - Gains (1/2)

Gains

☉ 목표범주 1(true)에 속하는 개체들이 각 등급에 얼마나 분포하고 있는지를 나타냄

➡ $\frac{\text{해당 등급에서 목표변수의 특정 범주 빈도}}{\text{전체에서 목표변수의 특정 범주 빈도}} \times 100$

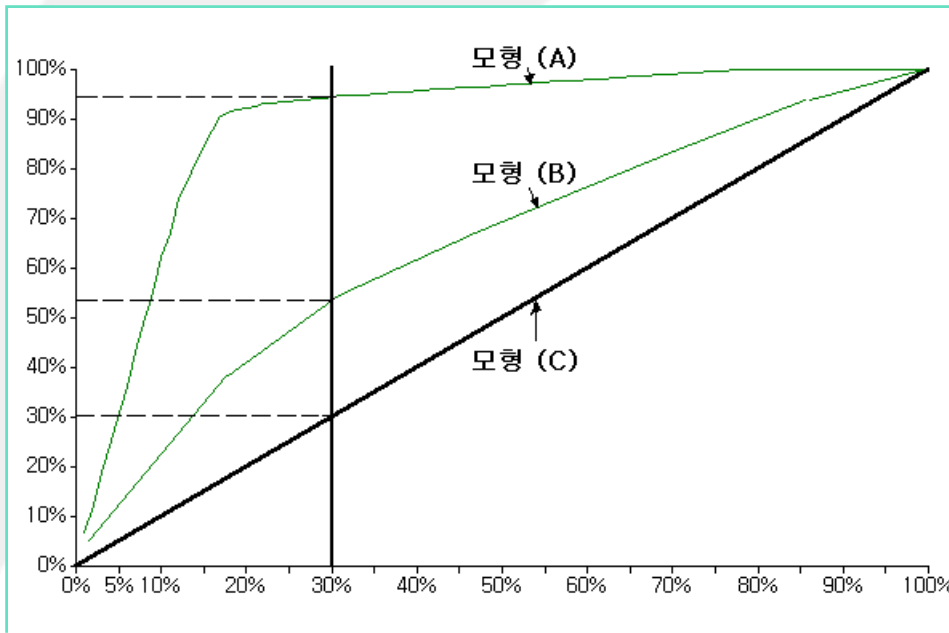


등급	비누적				누적			
	빈도			반응률	빈도			반응률
	합계	Y=1	Y=0	Gain	합계	Y=1	Y=0	Gain
1	200	174	26	174/381=45.6	200	174	26	174/381=45.6
2	200	110	90	110/381=28.8	400	284	116	284/381=74.5
3	200	38	162	38/381= 9.9	600	322	278	322/381=84.5
4	200	14	186	14/381= 3.6	800	336	464	336/381=88.1
5	200	11	189	11/381= 2.8	1000	347	653	347/381=91.0
6	200	10	190	10/381= 2.6	1200	357	843	357/381=93.7
7	200	7	193	7/381= 1.8	1400	364	1036	364/381=95.5
8	200	10	190	10/381= 2.6	1600	374	1226	374/381=98.1
9	200	3	197	3/381= 0.7	1800	377	1423	377/381=98.9
10	200	4	196	4/381= 1.0	2000	381	1619	381/381=100
전체	전체 반응률 =381/2000=19%							

> 모형 평가 방법 - Gains (2/2)

Gains Chart

- ④ 해당 등급에 따라 계산된 Gain값을 연속적으로 연결한 도표
- ④ 차트에서 볼 수 있는 좌하에서 우상을 걸친 대각선은 모형비교의 기준선으로서, 모형성능이 나쁘면 나쁠수록 이 기준선에 가까워짐



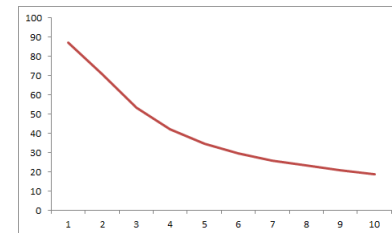
- ④ DM 발송에 대한 반응여부라는 목표변수의 '반응했음'이라는 범주에 대한 차트라고 가정해 보자
- ④ 서로 다른 방법으로 구축된 3개의 모형으로 동일한 수의 DM을 발송하는 경우, 얻어지는 반응률의 차이를 알 수 있음
- ④ 전체 관찰치 중 30%를 대상으로 DM을 발송하였을 때, 모형 (A)는 90%이상의 반응을 보임
- ④ 반면 모형 (B)는 50% 조금 넘는 반응을 보임
- ④ 따라서 분석자는 모형 (A)를 선택하게 됨

> 모형 평가 방법 - Response

Response

④ 각 등급에서 목표범주 1(true)의 비율을 나타냄

➡ $\frac{\text{해당 등급에서 목표변수의 특정 범주 빈도}}{\text{해당 등급에서 전체 빈도}} \times 100$



등급	비누적				누적			
	빈도			반응률	빈도			반응률
	합계	Y=1	Y=0	Response	합계	Y=1	Y=0	Response
1	200	174	26	174/200=87.0	200	174	26	174/200=87.0
2	200	110	90	110/200=55.0	400	284	116	284/400=71.0
3	200	38	162	38/200=19.0	600	322	278	322/600=53.6
4	200	14	186	14/200=7.0	800	336	464	336/800=42.0
5	200	11	189	11/200=5.5	1000	347	653	347/1000=34.7
6	200	10	190	10/200=5.0	1200	357	843	357/1200=29.7
7	200	7	193	7/200=3.5	1400	364	1036	364/1400=26.0
8	200	10	190	10/200=5.0	1600	374	1226	374/1600=23.3
9	200	3	197	3/200=1.5	1800	377	1423	377/1800=20.9
10	200	4	196	4/200=2.0	2000	381	1619	381/2000=19.0

> 모형 평가 방법 - Lift

Lift

- ② 전체 반응률에 비해 각 등급에서 반응률이 얼마나 높은지를 나타냄
- ② 상위 등급에서의 Lift가 매우 크고 하위 등급으로 갈수록 Lift가 감소하면 이는 모형의 예측력이 적절함을 의미함 등급에 관계없이 Lift에 별 차이가 없다면 이는 모형의 예측력이 좋지 않음을 나타냄

해당 등급에서 반응률(response)

전체 반응률

등급	비누적				누적			
	빈도			반응률	빈도			반응률
	합계	Y=1	Y=0	Lift	합계	Y=1	Y=0	Lift
1	200	174	26	870/190=4.57	200	174	26	870/190=4.57
2	200	110	90	550/190=2.89	400	284	116	710/190=3.73
3	200	38	162	190/190=1.00	600	322	278	536/190=2.82
4	200	14	186	70/190=0.36	800	336	464	420/190=2.21
5	200	11	189	55/190=0.28	1000	347	653	347/190=1.82
6	200	10	190	50/190=0.26	1200	357	843	297/190=1.56
7	200	7	193	35/190=0.18	1400	364	1036	260/190=1.36
8	200	10	190	50/190=0.26	1600	374	1226	233/190=1.23
9	200	3	197	15/190=0.07	1800	377	1423	209/190=1.10
10	200	4	196	20/190=0.10	2000	381	1619	190/190=1.00
전체	전체 반응률 = 381/2000 = 19%							