

Provider: Tom Lee, 6 Sigma Leader at www.eKnew.com

기초통계 (10)

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{(\sum y^2) - (\sum y)^2}}$$

상관분석 (Correlation Analysis) 두번째 시간

오늘은 상관계수 (1에 가까우면 서로 관련이 있음 / -1에 가까우면 서로 반비례함 / 0에 가까우면 서로 관계가 없음)을 구하는 공식에 대하여 알아보도록 하겠다. 먼저 위의 공식을 사용하려면, 연속되는 x 값 및 y 값을 필요로 한다.

회원가입 (x) (단위: 1000명)	1	3	4	6	8	9	11	14
마케팅비용 (y) (단위: 1000만원)	1	2	4	4	5	7	8	9

	x	y	xy	x ²	y ²
	1	1	1	1	1
	3	2	6	9	4
	4	4	16	16	16
	6	4	24	36	16
	8	5	40	64	25
	9	7	63	81	49
	11	8	88	121	64
	14	9	126	196	81
합	56	40	364	524	256

r = 0.977

‘합’이라고 되어진 부분들을 미리 구해놓는다면 공식에 쉽게 대입이 가능하겠다.

위의 x (회원가입현황) 및 y (마케팅비용)의 변화를 보자. 위의 빨간색 원을 보면, 보다 많은 마케팅비용이 투입된 기간동안 회원가입 효과가 큼을 쉽게 알 수 있다. 우리가 5 가지 마케팅 활동을 행한다고 가정해보자. TV 광고가 가장 효과적일까? 아니면, E-mail Marketing? 위의 실험을 통해, 우리의 Business 에 가장 효과적이고, 적합한 Marketing Activities 가 무엇인지 확인하고자 한다면, 일정 기간동안 단 한가지씩의 마케팅 활동을 통한 결과 (회원가입현황)의 결과를 Correlation 분석을 통해 서로의 상관계수를 비교해보는 것도 하나의 방법이겠다. 위의 가정에 대한 상관계수는 0.977 이 나왔다. 분명히 정비례 관계가 있음을 나타내지만, 또 다른 마케팅 활동을 위처럼 시험하여 0.999 가 나왔다면, 이것은 보다 효과적인 Marketing 수단으로서 검토대상이 될 수 있음을 의미한다. 6 Sigma 에서 상관분석을 이용하는 가장 큰 목적은 잘못된 프로세스상에서 그 주된 원인을 찾기위함이다. 무슨 문제든 원인은 있게 마련이며, 근본 원인을 찾는 것이 무엇보다도 중요하기에 이러한 사소하게 느껴지는 계수 (r 값)이 그렇게

공지사항

다음시간에는 회귀분석에 대하여 알아보기로 하겠습니다. 회귀분석이란 한 변수의 값을 알면 또 다른 변수의 값에 대한 정보를 알려주는 툴로 역시 통계에선 기본적인 사항이며, 6 Sigma 활동상에서도 반드시 알고 넘어가야 할 분야이다.



Newsletter Status

- 6Sigma 개요
- 통계
 - 통계란?
 - Sample
 - Descriptive Data
 - 히스토그램
 - 런 차트
 - 중심경향치 Data
 - 중심분산치 Data
 - 정규분포
 - 비정규분포
 - Attribute Data 의 Sample Size 구하기
 - Variable Data 의 Sample Size 구하기
 - 가설검증
 - 상관분석
 - 회귀분석
 - 다중회귀분석
 - ANOVA
 - Control Chart
- DMAIC
 - Define
 - Measure
 - Analyze
 - Improve
 - Control
- DFSS
 - Define
 - Measure
 - Analyze
 - Design
 - Verify

You are Here!



단어장

DMAIC Process

Define(정의)-Measure(측정)-Analyze(분석)-Improve(개선)-Control(통제)의 기본 5 단계로 구성된 6 Sigma 의 가장 유명한 방법론이다. 문제가 있다면, 그 문제를 정의하고, 데이터를 수집하여 분석하여 개선할 수 있는 부분이 무엇인가를 찾아 개선하고, 개선된 요소가 계속 다시 나빠지지 않게 돌보는 과정이라 보면 되겠다. 6 Sigma 활동이 어렵지만은 않다! 분명히...