명제

- $X, Y is \in dependent \Rightarrow Corr(X,Y) = 0$
- $Corr(X,Y) = \pm 1 \Leftrightarrow y = a + bx$

먼저 상관계수의 범위에 대한 성질을 증명하고자 한다. 새로운 확률 변수 U와 V를 잡아준다.

$U = \frac{X-E(X)}{\sqrt{Var(X)}}, V = \frac{Y-E(X)}{\sqrt{Var(Y)}}$

X,Y의 정규화 꼴로 생긴 변수 U,V는 어떠한 것을 의미하는가? 기댓값 E(U)와 E(V)를 구하면 0이 나온다.

U,V의 공분산에 대한 식은 다음과 같다.

$$Cov(U,V) = E(UV) - E(U)E(V)$$
$$= E(\frac{(X-E(X))}{Var(X)} \cdot \frac{(Y-E(X))}{Var(Y)}) - E(U)E(V)$$

$$\therefore Cov(U,V) = E(\frac{(X-E(X))}{Var(X)} \cdot \frac{(Y-E(X))}{Var(Y)}) \ \because E(U), E(V) = 0$$

$$= \frac{E((X-E(X))(Y-E(Y)))}{Var(X)Var(Y)} = \frac{Cov(X,Y)}{Var(X)Var(Y)} = Corr(X,Y)$$

$$\therefore Cov(U,V) = Corr(X,Y)$$

분산에 관한 성질이지만 서로 다른 두 확률변수의 분산에 대해서 다음의 식이 성립한다.

$$Var(U+V) = Var(U) + Var(V) + 2Cor(U,V)$$

여기서 다시 U와 V의 성질이 인다. 위에서 언급한 기댓값의 성질과 분산의 정의를 씀ㄴ 아래와 같이 분산의 값이 상수로 나온다.

$$Var(U) = Var(\frac{X-E(X)}{\sqrt{Var(X)}}) = E[(\frac{X-E(X)}{\sqrt{Var(X)}})^2] - (E(U)^2) = E[\frac{(X-E(X))^2}{Var(X)}]$$
$$= \frac{E[(X-E(X))^2]}{Var(X)} = \frac{Var(X)}{Var(X)} = 1$$

V에 대해서도 마찬가지 이다.

분산에 대한 성질이다. 분산은 평균과의 제곱 차이기에 언제나 0이상이다. 그에 대한 식은 다음과 같다.

$$0 \le Var(U+V) = Var(U) + Var(V) + 2Cov(U,V) = 1 + 1 + 2Cov(U,V)$$
$$= 2 + 2Cov(U,V)$$

$$\therefore \quad 0 \le 2 + 2Cov(U,V)$$

$$\therefore \quad -1 \le Cov(U,V) = Corr(X,Y) \ \cdots \text{식} 1$$

반대 방향에 대해서도 정리해보자.

$$0 \le Var(U-V) = Var(U) + Var(V) - 2Cov(U,V) = 1 + 1 - 2Cov(U,V)$$
$$= 2 - 2Cov(U,V)$$

$$\therefore \quad 0 \le 2 - 2Cov(U,V)$$

$$\therefore \quad 1 \ge Cov(U,V) = Corr(X,Y) \ \cdots \text{식} 2$$

식1과 식2를 결합하면 $-1 \le Corr(X, Y) \le 1$ 에 대한 증명이 완성된다.