Historia Mathematica Vol. 15, No. 2, 2002, 135-140

확률사상들의 독립성과 배반성 및 확률변수들의 독립성과 무상관성에 대하여

한양대학교 수학과 **장인홍**
서경대학교 수리정보통계학부 **박태룡**

Abstract

In this paper, we explain independence and mutually exclusiveness of events. Also we explain independence and uncorrelation of random variables. Finally we compare with differences as examples.

0. 서론

기초통계학과 확률론을 처음 배우는 학생들은 서로 배반사상과 독립사상 및 확률변수들의 독립성과 무상관성에 대하여 흔히 혼동을 일으키기 쉽다. 본 논문은 학생들이 일반적으로 일으키기 쉬운 혼동을 조금이나마 해소하는데 도움을 주기 위해 간단한 예들을 중심으로 전개해 나가고자 한다.

우선, 본 논문에서 사용되는 통계적 용어들에 대한 정의를 주기로 하자.

정의 1. 두 개의 확률사상 A와 B에 대하여 $A \cap B = \phi$(공사상)이면 A와 B는 배반(exclusiveness)이라 한다.

정의 2. 유한 개의 확률 사상 $A_1, A_2, \cdots, A_n$이 주어졌을 때 $A_i \cap A_j = \phi (i \neq j)$이면 $A_1, A_2, \cdots, A_n$은 서로 배반 혹은 상호 배반(mutual exclusiveness)이라 한다.

정의 3. 두 개의 확률 사상 A와 B에 대하여 $P(A \cap B) = P(A)P(B)$ 이면, A와 B는 독립(independence)이라 한다. (위에서, 임의의 확률 사상 C에 대하여 $P(C)$는 C가 일어나는 확률을 나타낸다.)

정의 4. 유한 개의 확률 사상 $A_1, A_2, \cdots, A_n$이 주어졌을 때 다음과 같은 모든 부분집합을 생각하자.

$$\{i_1, i_2, \cdots, i_k\} \subset \{1, 2, \cdots, n\} \quad (2 \le k \le n)$$

각 부분집합에 대하여 다음이 성립할 때 $A_1, A_2, \cdots, A_n$은 상호 독립 혹은 서로 독립(mutual independence)이라 한다.

$$P(A_{i1} \cap A_{i2} \cap \cdots \cap A_{ik}) = P(A_{i1})P(A_{i2})\cdots P(A_{ik})$$

정의 5. 두 개의 확률변수 X와 Y의 공분산(covariance)은 다음과 같이 정의된다.

$$E[(X-E(X))(Y-E(Y))]$$

이를 $Cov(X, Y)$로 표시한다.

또한, X와 Y의 상관도(correlation)는 다음과 같이 정의된다.

$$\frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}$$

이를 $\rho(X, Y)$로 표시한다. 만약 $\rho(X, Y) = 0$ 이면 X와 Y는 무상관인 확률변수들(uncorrelated random variables)이라 한다. (단, 임의의 확률변수 Z에 대하여 $E[Z]$와 $Var(Z)$는 각각 Z의 평균(mean)과 분산(variance)을 나타낸다.)

정의 6. 유한 개의 확률변수 $X_1, X_2, \cdots, X_n$에 대하여 $X_1, X_2, \cdots, X_n$의 결합확률함수나 확률밀도함수(joint probability function or probability density function)가 각각의 주변확률함수나 확률밀도함수(marginal probability function or probability density function)의 곱으로 나타내어지면 n개의 확률변수들은 (통계적)독립((statistical) independence)이라 한다. 즉, 다음이 성립할 때, $X_1, X_2, \cdots, X_n$은 (통계적)독립이라 한다.

$$f_{X_1, X_2, \cdots, X_n}(x_1, x_2, \cdots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

1. 본론

1.1. 독립사상과 배반사상

두 개 이상의 확률사상에서 각각의 양의 확률을 갖고 서로 배반이면 그들 중 하나의 사상이 일어났을 때 다른 사상들이 일어나는 것을 배제하기 때문에 명백히 상호 독립사상들이 아니다. 역으로 각각이 양의 확률을 갖고 상호 독립인 두 개 이상의 사상들은 서로 배반일

수 없다는 것을 쉽게 알 수 있다. 또한, 상호 독립이거나 혹은 상호 배반이 아닌 두 개 이상의 사상들에 대하여 두 개나 그 이상의 사상들의 모든 결합은 상호 독립이거나 혹은 서로 배반이 아니어야만 한다. 다음의 예는 쌍으로 독립(pairwise independence)이지만 상호 독립이 아닌 사상들의 존재에 대한 효과적인 설명으로서 사용될 수 있다.

예 1-1-1. 두 개의 공정한 동전이 던져질 때, 다음과 같은 세 개의 사상을 정의하자.

A = 첫 번째 동전에서 앞면이 나오는 사상
B = 두 번째 동전에서 앞면이 나오는 사상
C = 두 개의 동전 중에 정확히 한 개의 앞면이 나오는 사상

그러면 다음 두 가지 사실을 쉽게 알 수 있다.

$$P(A)=P(B)=P(C)=\frac{1}{2},$$

$$P(A\cap B)=P(A\cap C)=P(B\cap C)=\frac{1}{4}$$

이와 같이 A, B, C는 서로 배반이 아니고 쌍으로 독립이다. 그러나 A, B, C는 배반이다. 그러나 이와 같이 다음이 성립하기 때문에 A, B, C는 상호 독립이 아니다.

$$P(A\cap B\cap C)=P(\phi)=0\neq\frac{1}{8}=P(A)P(B)P(C)$$

한편으로, 다음은 사상 A, B, C는 독립이지만, 쌍으로 독립은 아닌 예로써 세 개의 사상 A, B, C는 서로 배반이 아님을 보여준다.

예 1-1-2. 두 개의 공정한 주사위를 던질 때, 다음과 같은 세 개의 사상을 정의하자.

A = 처음 주사위의 눈이 1, 2 또는 3이 나오는 사상
B = 처음 주사위의 눈이 3, 4 또는 5가 나오는 사상
C = 두 개의 주사위의 눈의 합이 9가 나오는 사상

그러면 $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{2}$, $P(C)=\frac{1}{9}$이고 $P(A\cap B\cap C)=\frac{1}{36}$이므로 다음이 성립한다.

$$P(A\cap B\cap C)=P(A)P(B)P(C)$$

또, 다음을 얻는다.

$$P(A \cap B) = \frac{1}{6},\ P(A \cap C) = \frac{1}{36},\ P(B \cap C) = \frac{1}{12}$$

그러므로 다음이 성립한다.

$$P(A \cap B) \neq P(A)P(B),\ P(B \cap C) \neq P(B)P(C),\ P(A \cap C) \neq P(A)P(C)$$

따라서 세 개의 사상 A, B, C는 상호 독립이 아니고 역시 서로 배반이 아니다.

1.2. 확률변수들의 독립성과 무상관성

확률변수들의 무상관성은 확률변수들이 독립이기 위한 필요충분조건은 아니다. 그러나 학생들은 예외의 경우인 이 변량 정규분포(bivariate normal distribution) 때문에 두 개의 개념을 동일하게 생각하는 경향이 있다.

두 개의 확률변수 X와 Y가 같은 확률분포를 따르고 있을 때, 다음과 같이 두 개의 새로운 확률변수 U와 V를 정의하자.

$$U = X - Y,\quad V = X + Y$$

직감적으로 U와 V는 독립이라고 생각될 수도 있다. 그러나 $E(U) = 0$ 이기 때문에 다음을 얻는다.

$$\begin{aligned} COV(U, V) &= E[(X-Y)(X+Y)] \\ &= E[(X^2 - Y^2)] \\ &= 0 \end{aligned}$$

이와 같이 U와 V는 무상관인 확률변수들이지만 독립일 수도 있고 독립이 아닐 수도 있다.

다음 예는 X와 Y는 독립이지만 U와 V는 독립이 아님을 보여준다.

예 1-2-1. 두 개의 공정한 주사위를 던질 때 확률변수 X와 Y를 다음과 같이 정의하자.

X = 첫 번째 주사위에서 나타나는 눈의 수
Y = 두 번째 주사위에서 나타나는 눈의 수

V가 취할 수 있는 값의 범위는 U의 값에 영향을 받는다.

예를 들면, $P(V=4 \mid U=3) = 0$이지만, $P(V=4) = \frac{3}{36}$ 이다.

이와 같이 X와 Y가 독립이지만 U와 V는 독립이 아니다.

한편, 다음 예는 X와 Y가 독립이 아니고, 같은 분포를 따르는 경우에 U와 V는 분명히 무상관이나 U와 V는 독립일 수도 있다는 것을 설명해 준다.

예 1-2-2. X와 Y가 이 변량 정규분포를 따르고 그들의 결합확률밀도함수가 다음과 같다고 하자. (단, $\rho = COV(x, y)$)

$$f_{X,Y}(x,y) = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp[-\frac{1}{2\pi(1-\rho^2)}(x^2-2xy\rho+y^2)]$$

이러한 경우에 U와 V의 결합확률밀도함수는 다음과 같다.

$$g_{U,V}(u,v) = \frac{1}{\sqrt{2\pi}\sqrt{2(1-\rho)}} \exp[-\frac{u^2}{4(1-\rho)}] \times \frac{1}{\sqrt{2\pi}\sqrt{2(1+\rho)}} \exp[-\frac{v^2}{4(1+\rho)}]$$

이와 같이 U와 V는 독립이고 각각이 평균 0 와 분산 $2(1-\rho)$ 와 $2(1+\rho)$ 를 갖는 일변량 정규분포(univariate normal distribution)를 따른다. 이 사실은 X와 Y가 같은 분포를 따를 때 U와 V는 독립일 수도 있다는 역설적인 예로서 사용될 수 있다.

2. 결론

확률 사상들의 서로 배반성과 완전 독립성은 완전히 다른 개념이라는 것을 1.1절에서 설명하였다. 또한, 확률변수 X와 Y의 무상관성이 X와 Y가 독립이기 위한 필요충분조건은 아니라는 사실은 1.2절에서 예를 들어 설명하였다. 그러나 X와 Y가 이 변량 정규분포를 따를 때, X와 Y의 무상관성은 X와 Y가 독립이기 위한 필요충분조건으로 잘 알려져 있다. 이 사실은 X와 Y가 각각 베르누이(Bernoulli) 확률분포를 따를 때도 성립한다는 것을 다음 예에서 볼 수 있다.

예 2-1-1. 확률변수 X가 취할 수 있는 값이 x_1, x_2이고 각각을 취할 확률은 P_1과 $1-P_1$이라 하자. 확률변수 Y가 취할 수 있는 값이 y_1, y_2라 하고 각각을 취할 확률은 P_2과 $1-P_2$라 하자. 그러면, X와 Y의 결합확률분포는 다음과 같이 쓰여진다.

$$P_{ij} = P(X=x_i, Y=y_j),\ \ i=1,2\ ;\ j=1,2$$

네 개의 값 x_1, x_2, y_1, y_2 중에 하나 (말하자면 x_1)를 0이라고 가정해도 일반성을 잃지 않으면서 동시에 증명을 간단하게 할 수 있다.

$\rho(X, Y) = 0$가 주어졌을 때 $COV(X, Y) = 0$ 이다.

이와 같이, 다음이 성립한다.

$$COV(X, Y) = y_1[P_{11}-(1-P_1)P_2]+y_2[P_{22}-(1-P_1)(1-P_2)]=0$$

$P_{22}+P_{21}=1-P_1$ 이라는 사실을 위의 식에 대입하면, 다음을 얻는다.

$$y_1[1-P_1-P_{22}-(1-P_1)P_2]+y_2[P_{22}-(1-P_1)(1-P_2)]=0$$

즉,

$$y_1[(1-P_1)(1-P_2)-P_{22}]+y_2[P_{22}-(1-P_1)(1-P_2)]=0$$

따라서 다음이 성립한다.

$$(y_2-y_1)[P_{22}-(1-P_1)(1-P_2)]=0$$

$y_1 \neq y_2$ 이기 때문에 $P_{22}=(1-P_1)(1-P_2)$ 이어야만 한다. 다른 세 가지 경우 $P_{11}=P_1P_2$, $P_{12}=P_1(1-P_2)$, 그리고 $P_{21}=(1-P_1)P_2$ 에서도 같은 과정을 통해 보여질 수 있다. 이와 같이 X와 Y는 독립이다.

참고 문헌

1. Casella, G., Berger, R. L., *Statistical Inference*. Brooks/Cole, 1990.
2. Hogg, R. V., Craig, A. T., *Introduction to Mathematical Statistics*, 5th ed., Macmillan, 1995.
3. Hogg, R. V., Tanis, E. A., *Probability and Statistical Inference*, 6th ed., Prentice Hall, 2001.
4. Mood, A. M., Graybill, R. A., Boes, D. C., *Introduction to the Theory of Statistics*, 3rd ed., Mcgraw-Hill, 1974.