|  |  |
| --- | --- |
| 교육 제목 | 확률과 통계 기초 |
| 교육 일시 | 2021.09.15 |
| 교육 장소 | 영우글로벌러닝 교육장 |
| **교육 내용** | |
| 오전 | ㅁ기초통계  연속형 자료  등간척도 -절대값 0없음, 곰셈법칙 적용x  비율척도 -절대값 0있음  범주형 자료  명목척도 -속성을 분류하는 척도 예) 남/녀, 국어/영어  순서척도 -속성간의 순위개념이 존재하는 척도 예) 상/중/하  모집단과 표본집단의 개념  모집단 -이해하고자 하는 집단전체  표본집단 - 모집단에서 추출된 집단  ex)표본집단이 연구에 참여한 고혈압 환자 100명이면 모집단은 국내 고혈압 환자 전체  범주형 자료  kxc 분할표, barplot, pie chart  연속형 자료  돗수분포표, 히스토그램, 상자수염 그림, 바이올린 그림  상자수염그림(boxplot) -수염부분은 mim과 max를 나타냄, 박스는 Q1,Q2,Q3으로 나뉘고 Q2는 median(0.5,중앙값)  그래프는 자료를 한눈에 파악 가능한 장점이 있지만 작성자의 주관에 따라 판단에 오류를 줄 수 있어 수치자료를 근거로서 같이 써야함  수치를 통한 연속형 자료요약  중심위치의 측도  표본평균 - 자료의 이상치에 영향을 많이 받음  중앙값 - 자료의 이상치에 영향을 적게 받음 |
| 오후 | ㅁ기초통계  퍼짐의 정도  분산 -S^2  표준편차(sd) -루트S^2  범위 -최댓값 - 최솟값  사분위수범위 -3사분위수 - 1사분위수  상관분석  두 연속형 변수간 선형적 연관관계가 있는지 분석하는 통계적 방법  연관성을 나타내는 척도이며 인과성을 설명하는 것은 아님  피어슨 상관계수  r값이 +일 경우 양의 상관관계 -일 경우 음의 상관 관계를 의미  r값의 범위는 -1~1까지 분포 1에 가까울 수록 높은 상관성  정규분포를 따르는 경우 사용  스피어만 상관계수  두 변수간 상관관계를 나타내는 비모수적 방법  정규분포를 따르지 않는 경우 사용  r값의 범위는 -1~1까지 분포 1에 가까울 수록 높은 상관성  ㅁ확률  사건의 확률  표본공간 - 한 실험에서 나올 수 있는 결과들의 모임  예) 동전 던지기 {앞, 뒤}, 주사위 던지기 {1, 2, 3, 4, 5, 6}  근원사건 - 표본공간을 구성하는 개개의 결과  사건 - 표본공간의 부분집합으로 근원사건들의 집합  조건부확률 - 특정 조건 하에 사건이 일어날 확률로 어떠한 조건 하인가에 따라 사건이 일어날 확률이 달라짐  조건부확률이 자주 나오고 중요함  독립사건 - 사건 a와 b가 있을때, 각각의 사건이 일어날 확률이 다른 사건에 영향을 주지 않는 경우 예) 동전 던지기, 주사위 |