5. 주요 확률 분포함수

확률 분포함수를 배우는 이유

통계 분석 과정에서는 수집된 데이터들이 특정 확률 분포함수를 따른다는 가정을 통해 이뤄지는 경우가 대부분

확률 분포함수를 알게 되면 구하고자 하는 사건에 대한 확률과 각종 정보의 유추가 가능해 짐

확률 분포함수의 주요 학습 내용

종류	내용
정의	확률 변수가 어떤 상황에 대하여 정의하고 있는지
범위	확률 변수가 취할 수 있는 값의 범위는 어떻게 되는지
형태	확률 분포 함수가 어떤 식으로 이루어져 있는지
요약값	확률 변수의 기댓값과 분산은 어떻게 계산되는지
모수	확률 분포함수의 모양, 형태 및 특성을 결정하는 값

균등 분포 (Uniform Distribution)

정의	서로 다른 n 개의 이산점에서 발생하는 확률이 1/n 로 동일한 확률변수
표현	$X \sim Uniform\ (n)$
형태	$f(x) = P(X = x) = \frac{1}{n}, \qquad x = 1, 2, 3, \ldots, n$
모수	n (변수 값의 개수)
요약값	$E(X) = \frac{n+1}{2}, \qquad V(X) = \frac{(n-1)(n+1)}{12}$

베르누이 분포 (Bernoulli Distribution)

정의	사건의 결과가 2가지로만 나타나는 시행 또는 실험 (성공 / 실패)
표현	$X \sim Bernouii\ (p)$
형태	$f(x) = P(X = x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$
모수	p (성공 확률)
요약값	$E(X) = p, \qquad V(X) = p(1-p)$

이항 분포 (Binomial Distribution)

정의	베르누이 시행을 n번 독립적으로 반복 시행한 경우에 나타나는 성공 횟수에 대해 정의
표현	$X \sim B\ (n, p)$
형태	$f(x) = P(X = x) = nCx\ p^x(1-p)^{1-x}, \qquad x = 0, 1, 2, \ldots, n$
모수	n (시행 횟수), p (성공 확률)
요약값	$E(X) = np, \qquad V(X) = np(1-p)$

포아송 분포 (Poisson Distribution)

정의	단위 구간 내에서 특정 사건이 발생하는 횟수에 대하여 정의하는 확률 변수
표현	$X \sim Poisson\ (\mu)$
형태	$f(x) = P(X = x) = \frac{\mu^x}{x!} e^{-\mu}, \qquad x = 0, 1, 2, \ldots$
모수	μ (단위 구간 내에서 특정 사건이 평균적으로 발생하는 횟수)
요약값	$E(X) = \mu, \qquad V(X) = \mu$

균등 분포 (Uniform Distribution)

정의	특정 구간에 속할 확률이 동일하도록 정의되는 확률 변수
표현	$X \sim Uniform\ (a, b)$
형태	$f(x) = \frac{1}{b-a}, \quad a \le x \le b$
모수	a (구간의 시작점), b (구간의 끝점)
요약값	$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}$

정규 분포 (Normal Distribution)

정의	평균을 중심으로 좌우 대칭인 종모양 (Bell Shape)을 가지는 연속형 확률 변수
표현	$X \sim N\ (\mu, \sigma^2)$
형태	$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty \leq x \leq \infty$
모수	μ (평균), σ^2(분산)
요약값	$E(X) = \mu, \qquad V(X) = \sigma^2$

표준 정규 분포 (Standard Normal Distribution)

정의	정규 분포를 표준화하여 하나의 분포 함수로 표현하고자 한 것
이유	정규분포에서 평균과 분산이 가질 수 있는 값이 무한개인데 이를 비교하고 싶어 생성
표현	$X \sim N\ (0,1)$
형태	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty \le x \le \infty$
표준화	$z = \frac{X - \mu}{\sigma}$
요약값	$E(X) = 0, \quad V(X) = 1$

T 분포 (Poisson Distribution)

정의	0을 중심으로 좌우 대칭인 종 모양의 분포이며, 정규 분포 모수 중 분산을 모를 때 활용
표현	$T \sim t\ (n)$
형태	$f(x) = \dfrac{Z}{\sqrt{V/\nu}}, \qquad -\infty \ \leq\ x \leq\ \infty$
모수	n(자유도) , 자유도가 커질수록 표준 정규분포에 가까워짐 (분산이 1에 가까워짐)
요약값	$E(X) = 0, \qquad V(X) = \dfrac{n}{n-2} > V(Z) = 1$