다변수 확률변수

정의

- 개요

1.여러 개의 확률 변수 $x_1, x_2, \dots, x_n$이 존재할 때,

2.확률벡터의 집합을 $\begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$으로 나타낼 수 있다.

- 다변량 확률벡터의 cdf와 pdf

1. n개의 확률변수의 벡터 $\vec{x}$가 이산형일경우 CDF는

- $F_{x_1,\dots,x_n}(x_1, \dots, x_n) = \sum \cdots \sum p[w_1, w_{2,} \cdots, w_{n,}]$

2. n개의 확률변수의 벡터 $\vec{x}$가 연속형일경우 CDF는

- $F_{x_1,\dots,x_n}(x_1, \dots, x_n) = \int \cdots \int_{-\infty}^{\infty} f[w_1, w_{2,} \cdots, w_{n,}]\, dx$

3. N개의 확률변수 벡터 $\vec{x}$에 대하여 PDF는

- $\frac{\partial F_{x_1,\dots,x_n}(x_1,\dots,x_2)}{\partial x_1,\dots,x_n} = f_{x_1,\dots,x_n}(x_1, \dots, x_n)$

정의

- 다변량 확률벡터의 기댓값

1. $\vec{x}= \begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix}$ 이고, Y = $u(x_1, \dots, x_n)$ 일 때

1) $\vec{x}$가 연속형일 경우 : $\int \cdots \int_{-\infty}^{\infty} u(x_1, \dots, x_n) \cdot f\left[w_1, w_{2,} \cdots, w_{n,}\right] dx$

2) $\vec{x}$가 이산형일 경우 : $\sum \cdots \sum u(x_1, \dots, x_n) \cdot p\left[w_1, w_{2,} \cdots, w_{n,}\right]$

3) 만약 Y가 벡터 $\vec{y} = \begin{bmatrix} y_1 \\ \cdots \\ y_n \end{bmatrix}$일 경우, $E[\sum k_j y_j]= \sum k_j E[y_j]$ 이다.

정의

- 다변량 확률변수의 주변 pdf

1) $\vec{x} = \begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix}$ 이고, 이 확률벡터가 결합 pdf를 가질 때,

(1) 조건부 $pdf\ f_{x_1}(x_1) = \int \cdots \int_{-\infty}^{\infty} f[x_1, x_2 \cdots, x_n]\, dx_2\, dx_3 \cdots dx_n$

- 즉, 해당 변수를 제외한 나머지 변수로 적분을 밀어 올린다.

(2) $\overrightarrow{x_k}$ = 를 $\vec{x}$와 선형종속 관계에 있는 벡터라고 가정하고,

- $\overrightarrow{x_k} = \begin{bmatrix} x_1 \\ x_3 \\ x_7 \end{bmatrix}$ 라고 한다면, 그 다변량 주변확률밀도함수는

- $f_{x_{1,}x_3,x_7}\left(x_{1,}x_3, x_7\right) = \int \cdots \int_{-\infty}^{\infty} f[x_1, x_2 \cdots, x_n]\, dx_2\, dx_4\, dx_5\, dx_6\, dx_8 \cdots dx_n$

정의

- 다변량 확률변수의 조건부 pdf

1. 어떤 조건부 pdf가 $f_{x_{1,}x_{3,}x_{7}|x_{2,}x_{4}}(x_{1,}x_{3}, x_{7}|x_{2,}x_{4})$ 로 정의되었을 경우

전체 결합 분포의 pdf $f_{x_{1},\ldots,x_{n}}(x_{1}, \ldots, x_{n})$ 에 대하여

1) $f_{x_{1,}x_{3,}x_{7}|x_{2,}x_{4}}(x_{1,}x_{3}, x_{7}|x_{2,}x_{4}) = \int \cdots \int_{-\infty}^{\infty} \frac{f[x_{1,}x_{2}\cdots,x_{n}]}{f[x_{2,}x_{4}]} dx_{2}\, dx_{4}\, dx_{5}\, dx_{6}\, dx_{8} \cdots dx_{n}$ 이다.

정의

- 다변량 조건부 기댓값

1. 어떤 관심있는 통계량 Y = u($x_1, x_3, x_7 | x_2, x_4$)의 기댓값 E(Y)는

1) $\int \cdots \int_{-\infty}^{\infty} u(x_1, x_3, x_7) \frac{f[x_1, x_2 \cdots, x_n]}{f[x_2, x_4]} dx_2\, dx_4\, dx_5\, dx_6\, dx_8 \cdots dx_n$ 으로 표현한다.

(1) 즉, 관심 변수들로 이루어진 통계량 u(x_1, x_3, x_7)를 조건부 pdf로 적분한 것이다.

정의

- 확률변수 벡터 $\vec{x}$가 서로 독립일 경우

1. 결합 pdf는 각 pdf의 곱과 같다.

1) $f[x_1, x_2 \cdots, x_n] = \prod_{i=1}^{n} f(x_i)$

2. 결합 cdf는 각 cdf의 곱과 같다.

1) $\mathrm{F}[x_1, x_2 \cdots, x_n] = \prod_{i=1}^{n} F(x_i)$

3. 확률은 각 확률의 곱이다.

1) $\mathrm{P}[a_1 < x_1 < b_1, a_2 < x_2 < b_2 \cdots, a_n < x_n < b_{2n}] = \prod_{i=1}^{n} \mathrm{P}[a_i < x_i < b_i]$

4. 기댓값은 각 기댓값의 곱이다.

1) $\mathrm{E}[u(x_1), u(x_2) \cdots, u(x_n)] = \prod_{i=1}^{n} \mathrm{E}[u(x_n)]$

5. mgf는 각 mgf의 곱이다.

1) $\mathrm{M}[t_1, t_2, \cdots, t_n] = \prod_{i=1}^{n} \mathrm{M}[0, \cdots, t_i, \cdots, 0]$

정의

- 확률변수 벡터 $\vec{x}$가 서로 독립일 경우

2) 만약, 확률변수 $\vec{x}$가 모두 mgf를 가질 경우, 그 결합 mgf는

(1) $\mathrm{M}[t_1, t_2, \cdots, t_n] = \int \cdots \int_{-\infty}^{\infty} e^{t\vec{x}^T} f[x_1, x_2 \cdots, x_n]\, dx_1 \cdots dx_n$

3) 이 때, 조건부 mgf만 구하고 싶다면

(1) $\mathrm{M}[t_1, t_2, \cdots, t_n | x_2, x_4] = \int \cdots \int_{-\infty}^{\infty} e^{t\vec{x}^T} \frac{f[x_1, x_2 \cdots, x_n]}{f[x_2, x_4]} dx_1 \cdots dx_n$