변환

정의

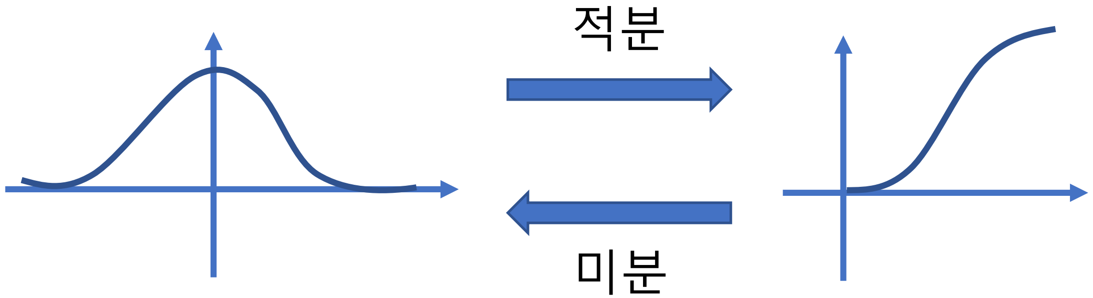
- CDF$f_x(x)$가 모든 포인트 x에 대해서 연속일 때, x를 연속형 확률변수라고 한다.

1) $p_x(x) = F_x(x) - F_x(x-)$라고 정의되므로,

모든 x에 대하여 $p_x(x)$ = 0이다.

(1) 즉, 연속확률변수는 도약점을 가지지 않기 때문에 그 지점에서만 정확한 확률값은 0에 수렴한다.

2) PDf와 CDF의 관계



정의

- 연속확률변수의 변환

1) X가 PDf $f_x(x)$ 를 가질 때,

(1) $\mathrm{y} = g_x(x)$ 라고 하고, 받침s_x 와 s_y는 1:1로 서로 대응되는 전단사 함수 관계라고 할 때

(2) $f_g(y) = P(Y \leqq y) = p\big(y \leqq g(X)\big) = P\big(x \leqq g^{-1}(y)\big) = F_x\big(g^{-1}(y)\big)$

2) 이 때, PDF $f_x(x)$ 는 CDF $\mathrm{F}(x)$ 의 미분이므로

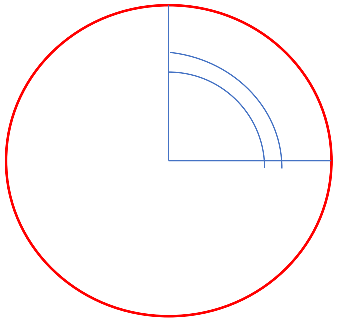
(1) $f_g(y) = \frac{F_x\left(g^{-1}(y)\right)}{\Delta y} = f_x\big(g^{-1}(y)\big) \cdot \frac{g^{-1}(y)}{\Delta y}$ (Chain Rule)

3) 정리하면, 역함수를 변수로 투입하여 구한 CDF를 역함수의 미지수로 미분한 Chain-Rule의 값이 변환 함수의 PDF가 된다.

(1) 이 때, 체인룰에 의해 튀어나온 $\frac{g^{-1}(y)}{\Delta y}$를 변환 야코비안이라고 한다.

예제

- R=1인 원 안에서 점을 무작위로 선택하는 실험이 있다. X를 원점에서 점 까지의 거리라고 할 때, 점이 $P\left(\frac{1}{4} < x < \frac{1}{2}\right)$에 속하는 확률을 구하시오



$$P(x \leq x) = \frac{\pi x^2}{\pi}$$

1) 이 때, CDF는 $F_x \begin{cases} 0 & x<0 \\ x^2 & x < 1 \\ 1 & x \geq 1 \end{cases}$

2) 이 때, PDF는 $f_x \begin{cases} 0 & x < 0 \\ 2x & 0 < x < 1 \\ 1 & x > 1 \end{cases}$

3) 점이 $\frac{1}{4} < x < \frac{1}{2}$ 에 속하는 확률은 $\int_{\frac{1}{4}\cdot}^{\frac{1}{2}} 2x\,dx = [x^2]_{\frac{1}{4}}^{\frac{1}{2}} = \frac{1}{4} - \frac{1}{10} = \frac{3}{16}$

예제

- Y = 2logx의 관계를 가지는 확률변수 X,Y가 있을 때

1) x의 PDf $\begin{cases} 1 & 0 < x < 1 \\ 0 & else \end{cases}$

2) $Y^{-1} \rightarrow -\frac{1}{2}Y = logx, x = e^{-\frac{1}{2}y}$

3) F_x 를 미분하여 pdf로 변환하면

(1) $\frac{F_x(e^{-\frac{1}{2}y})}{\Delta y} = \frac{F_x(e^{-\frac{1}{2}y})}{\Delta e^{-\frac{1}{2}y}} \cdot \frac{e^{-\frac{1}{2}y}}{\Delta y} = f_x(e^{-\frac{1}{2}y}) \cdot e^{-\frac{1}{2}y} \cdot (-\frac{1}{2})$

(2) 즉, $f_y(y) = f_x(e^{-\frac{1}{2}y}) \cdot (-\frac{1}{2})e^{-\frac{1}{2}y}$