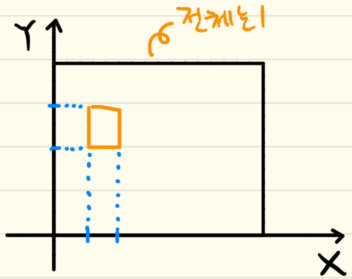
< 통계 2장 문제 >

12, 13, 19

2.33 Theorem. *Suppose that the range of X and Y is a (possibly infinite) rectangle. If $f(x,y) = g(x)h(y)$ for some functions g and h (not necessarily probability density functions) then X and Y are independent.* 12번

X, Y 가 independent 하다는 것은 if and only if

$\Rightarrow f_{X,Y}(x,y) = f_X(x) f_Y(y)$ → 근데 이건 $f_X(x)$, $f_Y(y)$가 pdf 일때



rectangle이고 $f(x,y)$가 $g(x)h(y)$로 표현된다면 X, Y가 independent 한가?

① rectangle의 의미는?
② $g(x)$, $h(y)$ 가 pdf가 아니어도 되나?

$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ 를 만족하면 X, Y가 independent 하다고 함.

↳ rectangle은 이를 만족하는것같은데

rectangle이니까 $f(x,y) = f_X(x) f_Y(y)$

g, h가 pdf가 아닌 예, $g(x) = 2 \cdot f_X(x)$
$h(y) = \frac{1}{2} \cdot f_Y(y)$

뭔가 엄밀하지 않군..

뒷장에 박사님 풀이!

2.33 Theorem. *Suppose that the range of X and Y is a (possibly infinite) rectangle. If $f(x,y) = g(x)h(y)$ for some functions g and h (not necessarily probability density functions) then X and Y are independent.*

sungbin 🤗 5:52 PM
1. 원래 Rectangle 조건은 필요없어요. Measure theory 안 쓰고 저걸 설명하려다보니 저자가 rectangle 이 필요하다고 쓴 것임

Borel set R.V.

$X \to A \to X$

여길보시오

2. P(X \in dx) 는 적분 안에 measure 를 표기하는 기호에요. 해석하자면 "x 에 따라 P_{X} 라는 분포 측도로 적분하라" 라는 뜻임요.

보통 학부 확률론에선 P(X \in dx) 대신 pdf 를 써서 f(x)dx 라고 쓰죠.

그래서 "x 에 따라 f(x) 를 곱해서 적분하라" 라는 뜻이 되는거

하지만 advanced probability theory 에선 pdf 대신 저런 distribution 기호를 써요

왜냐하면 -
1. probability density function 이 항상 존재하는게 아니고
2. pdf 를 쓰는 것보다 abstract 하게 의미전달이 가능해서

1에서 저자가 rectangle 이라고 쓴 건 X 랑 Y 가 좌표계로 각각 mapping 되는 확률변수를 고려하겠다는 의미.. 그래서 X 는 \mathcal{X} 로 가고 Y 는 \mathcal{Y} 로 가는 setting 에서 \sigma(\mathcal{X}) 랑 \sigma(\mathcal{Y}) 를 쓴거에요

$\in$

Suppose $f(x,y) = g(x)h(y)$ for some function g and h. Let $A \in \sigma(\mathcal{X})$ and $B \in \sigma(\mathcal{Y})$ be any Borel sets of $\mathcal{X}$ and $\mathcal{Y}$ respectively. If $B = \mathcal{Y}$, then

$$\mathbb{P}(X \in A) = \int_{A\times\mathcal{Y}} \mathbb{P}(X \in dx, Y \in dy)$$

왜 dx, dy 지? A, y 아닌가?

$$= \int_A \int_{\mathcal{Y}} f(x,y)dxdy = \int_A \int_{\mathcal{Y}} g(x)h(y)dxdy$$

$$= \left(\int_A g(x)dx\right) \times \left(\int_{\mathcal{Y}} h(y)dy\right)$$

분리!

Since $\int_{\mathcal{Y}} f(x,y)dy = f(x)$ holds,

$$\int_A f(x)dx = \left(\int_A g(x)dx\right) \times \left(\int_{\mathcal{Y}} h(y)dy\right)$$

Note that the above equality holds for every $A \in \sigma(\mathcal{X})$, one can see that

전체에대해 적분하면 →
$$1 = \int_{\mathcal{X}} f(x)dx = \left(\int_{\mathcal{X}} g(x)dx\right) \times \left(\int_{\mathcal{Y}} h(y)dy\right)$$

Hence

$$\int_{\mathcal{Y}} h(y)dy = \frac{1}{\int_{\mathcal{X}} g(x)dx}$$

Therefore

$$\mathbb{P}(X \in A) = \frac{\int_A g(x)dx}{\int_{\mathcal{X}} g(x)dx}$$

$\Leftrightarrow \int_A g(x)dx = P(X \in A) \cdot \int_{\mathcal{X}} g(x)dx$

Similarly,

$$\mathbb{P}(Y \in B) = \frac{\int_B h(y)dy}{\int_{\mathcal{Y}} h(y)dy}$$

Then observe that

$$\mathbb{P}(X \in A, Y \in B) = \int_{A\times B} \mathbb{P}(X \in dx, Y \in dy)$$

$$= \int_A \int_B f(x,y)dxdy \qquad (\because \text{ Definition of density function})$$

$$= \int_A \int_B g(x)h(y)dxdy$$

$$= \int_A g(x)\left[\int_B h(y)dy\right]dx$$

$$= \left(\int_A g(x)dx\right) \times \left(\int_B h(y)dy\right)$$

$$= \frac{\int_A g(x)dx}{\int_{\mathcal{X}} g(x)dx} \times \frac{\int_B h(y)dy}{\int_{\mathcal{Y}} h(y)dy} \times \int_{\mathcal{X}} g(x)dx \times \int_{\mathcal{Y}} h(y)dy$$

$$= \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \times \underbrace{\int_{\mathcal{X}} g(x)dx \times \int_{\mathcal{Y}} h(y)dy}_{=1}$$

where

$$1 = \mathbb{P}(X \in \mathcal{X}, Y \in \mathcal{Y}) = \left(\int_{\mathcal{X}} g(x)dx\right) \times \left(\int_{\mathcal{Y}} h(y)dy\right)$$

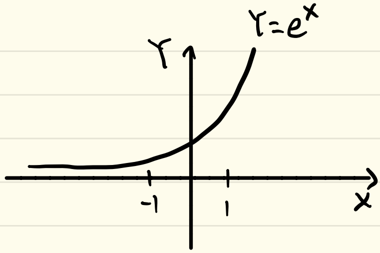
or rectangle이란 조건이 있어서 가능한건가?

Therefore

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

13. Let $X \sim N(0,1)$ and let $Y = e^X$.

(a) Find the PDF for Y. Plot it.



$Y > 0$

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \le y) = \mathbb{P}(r(X) \le y) \\ &= \mathbb{P}(\{x;\ r(x) \le y\}) \\ &= \int_{A_y} f_X(x)dx. \end{aligned}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\},$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2} \quad (\mu=0,\ \sigma=1)$$

적접안구하고 CDF로 그냥 정리

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(e^X \leq y) \\ &= P(x \leq \log y) \\ &= \int_{-\infty}^{\log y} f_X(x)\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\log y} e^{-x^2} dx = F_X(\log(y)) \end{aligned}$$

$$\begin{aligned} f_Y(y) = F_Y(y)' &= F_X(\log(y))' \\ &= F_X'(\log y) \times \frac{1}{y} \\ &= f_X(\log y) \times \frac{1}{y} \\ &= \frac{1}{y\sqrt{2\pi}} e^{-(\log y)^2}, \quad y > 0 \end{aligned}$$

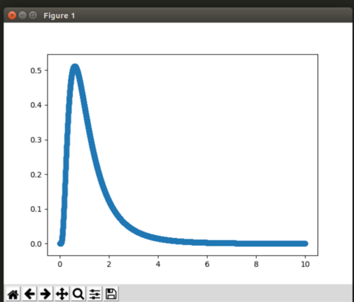
NormalExponential.py

```python
import numpy as np
import matplotlib.pyplot as plt

def pdf(x):
    return 1/(x*np.sqrt(2*np.pi)+1e-5)*np.exp(-(np.log(x)**2))

x = np.linspace(0.0, 10.0, num=10000)
y = pdf(x)

plt.scatter(x,y)
plt.show()
```



19.

When r is strictly monotone increasing or strictly monotone decreasing then r has an inverse $s = r^{-1}$ and in this case one can show that

$r(X) = Y$

$X = r^{-1}(Y)$

$$f_Y(y) = f_X(s(y)) \left| \frac{ds(y)}{dy} \right|. \quad \text{음?} \tag{2.12}$$

$$\begin{aligned} f_Y(y) &= P(Y = y) \\ &= P(r(x) = y) \\ &= p(\{x ; r(x) = y\}) \\ &= p(X \in r^{-1}(y)) \end{aligned}$$

$$\begin{aligned} F_Y(Y \leq y) &= P(X \leq r^{-1}(y)) \\ &= p(X \leq s(y)) \\ &= F_X(s(y)) \end{aligned}$$

$F_Y(y)' = f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(s(y)) \frac{ds(y)}{dy}$ 인데 $f_Y(y)$가 PDF이기 때문에 절대값을 씌움.