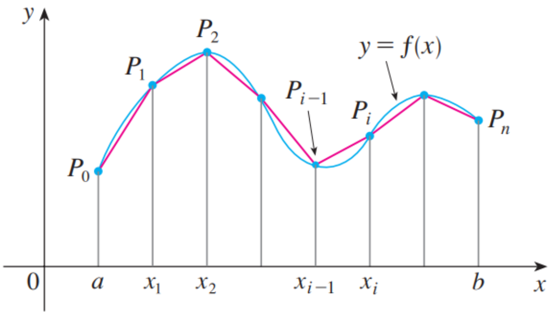
8. Further Applications of Integration

8.1 Arc Length

곡선의 길이를 알기란 쉽지 않다. 하지만 우리는 직선의 길이를 구하는 법은 알고 있다. 곡선을 아주 잘게 쪼개는 아래의 상황을 생각하여 보자.



점 x_i와 x_{i+1} 사이의 간격을 점점 줄여나가면 그 작은 구간에서는 곡선이 거의 직선과 동일하게 될 것이며, 길이도 거의 같아질 것이다. 즉 우리는 곡선의 길이 L에 대해 아래와 같은 생각을 할 수 있다.

$$L = \lim_{n\to\infty} \sum_{i=1}^{n} |P_{i-1}P_i|$$

이때 피타고라스 정리와 MVT에 의하여 $x_i^* \in (x_{i-1}, x_i)$가 존재하여

$$|P_{i-1}P_i| = \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2} = \sqrt{(x_i - x_{i-1})^2 + (x_i - x_{i-1})^2 f'(x_i^*)^2} = \sqrt{1 + f'(x_i^*)^2} \cdot (x_i - x_{i-1})$$

이므로, 위의 식에 대입하면 Riemann integral의 정의에 의하여

The Arc Length Formula
만약 f'이 $[a,b]$에서 continuous라면, 곡선 $y = f(x)$, $a \le x \le b$의 길이는

$$L = \int_a^b \sqrt{1 + [f'(x)]^2}dx$$

이다.

이때 f'의 연속성은 함수 $\sqrt{1+[f'(x)]^2}$가 연속할 수 있게 해줌으로써 Riemann integral이 존재하는 데 기여한다. 만약 곡선이 $x = g(y), c \le y \le d$ 꼴이 주어지더라도, 적절히 몇 개의 문자만 바꿔주면 된다. 이 경우 $g'(y)$가 continuous라면

$$L = \int_c^d \sqrt{1 + [g'(y)]^2}dy$$

가 성립함을 알 수 있다.

어떨 때는 특정 점 $P_0 = (a, f(a))$로부터 점 $Q = (x, f(x))$까지의 곡선의 길이를 x의 함수로 표현해야 하는 일도 있다. 그럴 때는 아래의 함수를 사용해줄 수 있다. 당연히 $f'(x)$는 continuous임을 가정한다.

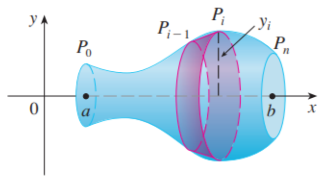
Arc length function

$$s(x) = \int_a^x \sqrt{1 + [f'(t)]^2}dt$$

그렇다면 FTC에 의해 $ds/dx = \sqrt{1+[f'(x)]^2}$임도 당연하다.

8.2 Area of a Surface of Revolution

곡선을 아주 잘게 나누어 작은 직선들의 합으로 보았었다. 회전체에서 우리는 그 표면적을 알기 위해, 회전체를 수없이 작은 원뿔대 여러 개의 합으로 볼 것이다.



(b) Approximating band

이때 붉은색 부분을 근사하는 원뿔대의 표면적은 간단한 계산을 통해

$$2\pi\frac{y_{i-1}+y_i}{2}|P_{i-1}P_i|$$

임을 확인해줄 수 있다. 또한 $x_i - x_{i-1}$가 매우 작아진다면, 함수 f의 연속성에 의하여 $y_{i-1} \approx y_i$이다. 그런데 앞선 계산에서

$$|P_{i-1}P_i| = \sqrt{1+f'(x_i^*)^2}\cdot(x_i - x_{i-1})$$

였다. 그러므로 우리가 원하는 회전체의 표면적은

$$S = \lim_{n\to\infty}\sum_{i=1}^{n} 2\pi\frac{y_{i-1}+y_i}{2}|P_{i-1}P_i| = \lim_{n\to\infty}\sum_{i=1}^{n} 2\pi f(x_i)\sqrt{1+[f'(x_i^*)]^2}\Delta x = \int_a^b 2\pi f(x)\sqrt{1+[f'(x)]^2}dx$$

임을 알 수 있다.

함수 f가 positive이고 continuous derivative를 가질 때, $y = f(x)$를 $a \le x \le b$에서 x축을 기준으로 회전시킨 회전체의 표면적은

$$S = \int_a^b 2\pi f(x)\sqrt{1+[f'(x)]^2}dx$$

이다.
한편 $x = g(y)$가 continuous derivative를 가지고 $x \ge 0$일 때, $c \le y \le d$에서 y축을 기준으로 회전시킨 회전체의 표면적은

$$\int_c^d 2\pi g(y)\sqrt{1+[g'(y)]^2}dy$$

이다.

더불어 arc length function $s(x)$에 대하여

$$\frac{ds}{dx} = \sqrt{1+[f'(x)]^2}$$

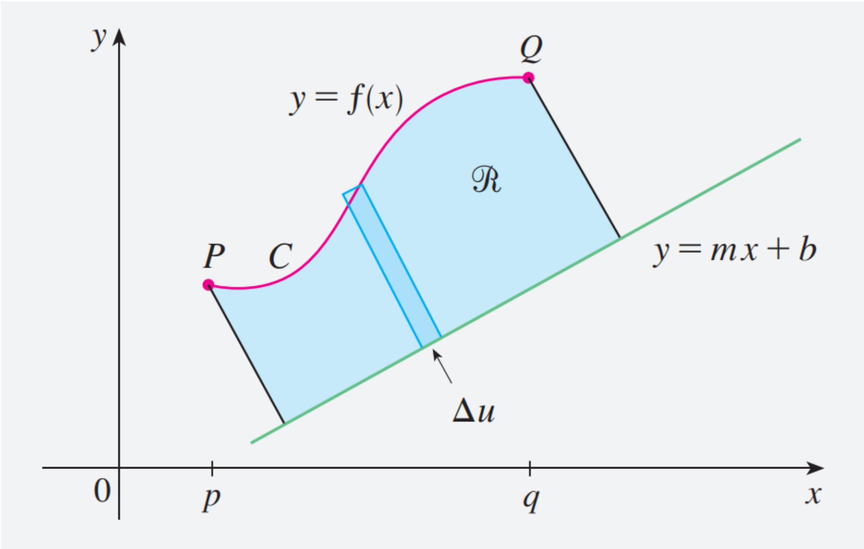
임을 보였었다. 그러면 surface에 관한 식에서 적분하는 변수를 x에서 s로 바꾸는 치환적분을 단행하면,

$$S = \int_{s(a)}^{s(b)} 2\pi y ds$$

임을 알 수 있다. 이때 $s(a) = 0$이고, $s(b)$는 계산해주거나, 간단히 주어질 수 있다. 단, y를 s에 대한 함수로 간단하게 표현할 수 없는 경우에는 이것이 계산을 더 복잡하게 만들 수도 있다. 일반적으로는 위의 박스에서 주어진 식을 더 선호한다.

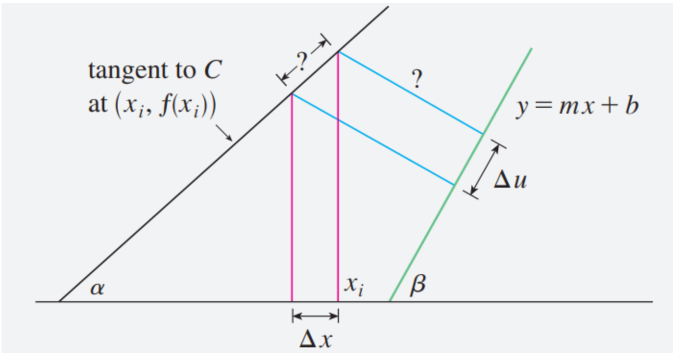
8.3 Rotating: Special Cases

회전체를 항상 x축이나 y축으로만 돌릴 수 있는 건 아니다. 시험에는 가끔씩 다른 직선에 대해 그래프를 회전시킬 때 얻는 회전체의 표면적을 구하는 문제도 등장한다.



여기서는 $f(x) > mx + b$인 상황에서 점 $P = (p, f(p))$와 $Q = (q, f(q))$ 사이의 $y = f(x)$ 그래프가 직선 $y = mx + b$를 기준으로 회전했을 때 얻어지는 회전체의 표면적을 구하고자 한다.

먼저 $\mathcal{R}$의 넓이부터 구해보면서 아이디어를 얻어 보자. 물론 항상 f의 도함수는 연속임을 가정하자.



점 $(x_i, f(x_i))$에서 $y = mx + b$에 내린 수선의 발이 $(u_i, mu_i + b)$라고 하여 보자. 그러면 두 점을 잇는 직선은 $y = mx + b$에 수직해야 하므로,

$$\frac{mu_i + b - f(x_i)}{u_i - x_i} = -\frac{1}{m}$$

을 얻을 수 있다. 이를 잘 정리해준다면 u_i에 대한 식을 얻을 수 있는데,

$$u_i = \frac{1}{1 + m^2} \cdot (mf(x_i) - mb + x_i)$$

이다. 수선의 길이는 피타고라스의 정리를 이용하여 자연스럽게

$$l_i = \sqrt{(u_i - x_i)^2 + (mu_i + b - f(x_i))^2} = \sqrt{1 + \frac{1}{m^2}}\,|u_i - x_i| = \frac{1}{\sqrt{1 + m^2}} \cdot (f(x_i) - mx_i - b)$$

임을 안다. 그러면 우리가 윗 그림의 진한 푸른색 직사각형의 넓이로써 전체 영역의 넓이를 근사한다면

$$\mathcal{R} = \lim_{n\to\infty}\sum_{i=1}^{n} l_i\Delta u = \lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{\sqrt{1+m^2}}\cdot(f(x_i)-mx_i-b)\Delta u$$

임을 얻는다. 이때 MVT에 의하여 $f(x_i)-f(x_{i-1}) = (x_i-x_{i-1})f'(x_i^*)$인 $x_i^*\in(x_{i-1},x_i)$가 존재하므로

$$\Delta u = \sqrt{1+m^2}\cdot(u_i-u_{i-1}) = \frac{1}{\sqrt{1+m^2}}(mf(x_i)-mf(x_{i-1}))+\frac{1}{\sqrt{1+m^2}}(x_i-x_{i-1}) = \frac{1}{\sqrt{1+m^2}}(1+mf'(x_i^*))\Delta x$$

이다. 따라서

$$\mathcal{R} = \lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{1+m^2}(f(x_i)-mx_i-b)(1+mf'(x_i^*))\Delta x = \frac{1}{1+m^2}\int_p^q[f(x)-mx-b][1+mf'(x)]dx$$

이다. 편의상 Δu라고 썼지만, 사실 이는 (u_i-u_{i-1})을 의미하는 것보다는 $(u_{i-1}, mu_{i-1}+b)$와 (u_i, mu_i) 사이의 거리를 의미하는 것이다. 이에 주의하자.

이러한 논리를 이용하면 회전체의 부피도 아주 쉽게 구할 수 있다. 부피는

$$V = \lim_{n\to\infty}\sum_{i=1}^{n}\pi l_i^2\Delta u = \lim_{n\to\infty}\sum_{i=1}^{n}\pi\cdot\frac{1}{1+m^2}(f(x_i)-mx_i-b)^2\cdot\frac{1}{\sqrt{1+m^2}}(1+mf'(x_i^*))\Delta x$$

임에 따라

$$V = \frac{\pi}{(1+m^2)^{3/2}}\int_p^q[f(x)-mx-b]^2[1+mf'(x)]dx$$

으로 주어진다.

마지막으로 우리의 목적인 표면적을 구할 차례이다. 아이디어는 정말정말 동일하다.

$$S = \lim_{n\to\infty}\sum_{i=1}^{n}2\pi\frac{l_{i-1}+l_i}{2}\sqrt{1+[f'(x_i^*)]^2}\Delta u$$

임을 떠올리기만 하면 된다. 2π는 상수이니 설명할 필요조차 없다. $\dfrac{l_{i-1}+l_i}{2}$는 원뿔대에서 윗면과 아랫면의 반지름의 평균이다. $\sqrt{1+[f'(x_i^*)]^2}$는 원뿔대의 모선부분 길이이다. 이제 이를 정리해주기만 하면 된다.

$$\begin{aligned}
S &= \lim_{n\to\infty}\sum_{i=1}^{n}2\pi\frac{l_{i-1}+l_i}{2}\sqrt{1+[f'(x_i^*)]^2}\Delta u\\
&= \lim_{n\to\infty}\sum_{i=1}^{n}2\pi l_i\sqrt{1+[f'(x_i^*)]^2}\Delta u\\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\frac{2\pi}{1+m^2}(f(x_i)-mx_i-b)\sqrt{1+[f'(x_i^*)]^2}(1+mf'(x_i^*))\Delta x\\
&= \frac{2\pi}{1+m^2}\int_p^q[f(x)-mx-b][1+mf'(x)]\sqrt{1+[f'(x)]^2}dx
\end{aligned}$$

를 얻을 수 있을 것이다.

이제 또 다른 특별한 경우를 이야기해보고자 한다. 가끔씩 어떤 곡선들은 한정된 구간 $[a,b]$에서 정의되지 않고 $[a,\infty]$와 같이 무한대 구간에서 정의되기도 한다. 다행히 이 경우 그냥 앞의 공식에서 b의 자리에 ∞를 넣어주면 된다. 물론 이 improper integral이 수렴할 때의 이야기이다.

$f(x) \geq 0$이거나 $f(x) \geq mx + b$인 경우가 아니라, 대소가 왔다갔다 하는 경우가 있을 수도 있다. 예를 들어 $y = \sin x$를 $[-\pi, \pi]$에서 회전시키는 경우 등을 생각해볼 수 있다. 어차피 중요한 것은 $f(x)$의 값이라기보다는 그래프와 회전시키는 직선 사이의 거리이므로, 우리는 $f(x)$ 대신 $|f(x)|$를 사용하면 된다. $y = mx+b$를 기준으로 회전시키는 경우에는 $|f(x) - mx - b|$를 사용하면 되겠다. 따라서

$$S = \int_a^b 2\pi|f(x)|\sqrt{1 + [f'(x)]^2}dx$$

을 이용하면 된다.

더욱 까다로운 경우도 있을 수 있다. 예를 들면 $y = \max\{\sin x, \cos x\}$의 그래프를 그린다면, x의 범위에 따라 $|f(x)|$의 자리에 들어가야 할 함수가 달라진다. 이런 경우 $|\sin x|$와 $|\cos x|$를 비교하여 각각이 더 큰 범위를 확인하고, 각 범위에서 따로 적분하여 더해주어야 하는 수고로움이 있다.

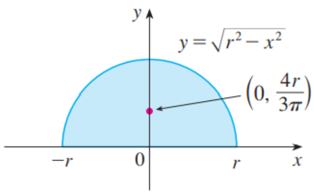
가장 짜증날 수 있는 경우는 도넛 모양의 회전체가 만들어질 때이다. $x^2 + (y-2)^2 = 1$으로 이루어지는 원을 x축을 기준으로 회전시킨다고 가정하여 보자. 그렇다면 이 도형을 어떻게 표현해야 할지 막막해진다. 이때는 주어진 도형을 여러 개의 함수 형태로 쪼개어 생각하여 보자. 그렇다면 도넛 모양을 기준으로 바깥쪽에 있는 부분은 $y = 2 + \sqrt{1-x^2}$으로, 안쪽에 있는 부분은 $y = 2 - \sqrt{1-x^2}$이다. 도넛의 표면적은 각각의 표면적을 더한 것이므로, 두 함수에 대해 S_1, S_2를 각각 구한 다음 이를 더해 전체의 표면적을 구해 주면 된다. 한 도형의 표면적을 계산하는 데에 두 번의 계산을 해야 한다는 점에서 부담이 크다.

8.4 Applications

곡선 $y = f(x)$와 $x = a$, $x = b$, x축으로 둘러싸인 부피의 **centroid** $(\bar{x}, \bar{y})$는 아래와 같이 구한다.

$$\bar{x} = \frac{1}{A}\int_a^b xf(x)dx, \bar{y} = \frac{1}{A}\int_a^b \frac{1}{2}[f(x)]^2dx$$

이때 A는 전체 영역의 넓이로, $A = \int_a^b f(x)dx$이다.



곡선 $y = f(x), y = g(x)$와 직선 $x = a, x = b$로 둘러싸인 부피의 centroid는

$$\bar{x} = \frac{1}{A}\int_a^b x[f(x) - g(x)]dx, \bar{y} = \frac{1}{A}\int_a^b \frac{1}{2}\{[f(x)]^2 - [g(x)]^2\}dx$$

으로 구한다. $A = \int_a^b [f(x) - g(x)]dx$이다.

증명은 수학보다는 물리에 더욱 가까우므로 생략하고자 한다. 대신 다음의 정리에 조금 더 집중하려 한다.

Theorem of Pappus
$\mathcal{R}$은 직선 l을 기준으로 한 쪽에만 있는 도형이다. 만약 이 도형이 직선 l을 기준으로 회전될 때, 얻어지는 도형의 부피는 $\mathcal{R}$의 넓이 A와 centroid가 이동하는 거리 d의 곱이다.

간단한 경우에 대해 증명하여 보자. $y = f(x)$와 $y = g(x)$, $x = a, x = b$ 사이에 둘러싸인 영역을 y축으로 회전시키려 한다. 그러면 부피는

$$\begin{aligned} V &= \int_a^b 2\pi x[f(x) - g(x)]dx \\ &= 2\pi \int_a^b x[f(x) - g(x)]dx \\ &= 2\pi(\bar{x}A) \\ &= Ad \end{aligned}$$

임을 알 수 있다. 이는 다른 도형에도 모두 적용될 수 있다.

8.5 Probability

random variables: 실험의 결과에 의해 서로 다른 값을 가지는 수치형 변수
indicator random variable: 특정 사건 A에 대하여, I가 아래와 같이 정의되면 이는 사건 A에 대한 indicator이다.

$$I = \begin{cases} 1 & (A \quad \text{happens}) \\ 0 & (A \quad \text{does not happen}) \end{cases}$$

discrete random variable: 가질 수 있는 값이 유한 개이거나, 셀 수 있는 무한 개인 random variable
continuous random variable: 가질 수 있는 값이 셀 수 없는 무한 개인 random variable

cumulative distribution function: random variable X에 대하여, 아래와 같이 정의되는 함수 F이다.

$$F(x) = P\{X \le x\}$$

만약 $X \sim F$처럼 쓴다면, random variable X가 cumulative distribution function으로 F를 가진다는 의미이다.

cumulative distribution function은 줄여서 CDF라고 쓰기도 한다. 또한 이를 이용하면 X가 특정 범위 안에 존재할 확률도 쉽게 구할 수 있다. 예를 들면, 아래가 성립한다.

$$P\{a < X \le b\} = P\{X \le b\} - P\{X \le a\} = F(b) - F(a)$$

한편 X가 특정 값을 가질 확률을 구하려면 어떻게 해야 할까? 먼저, discrete일 때는 아래와 같은 표기를 사용가능하다.

discrete random variable X에 대해, **probability mass function** $p(a)$는

$$p(a) = P\{X = a\}$$

로 정의한다. 이때 X가 $x_1, x_2, \cdots$의 값을 가진다면

$$p(x_i) > 0, \quad p(x) = 0 \quad (x \neq x_i)$$

이다.

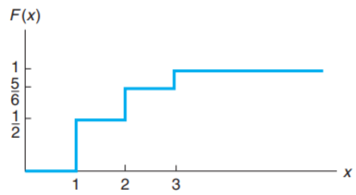
또 전체 확률은 항상 1이어야 하기에, 아래 등식 또한 성립한다.

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

또 정의에 의하여 cumulative distribution function은

$$F(a) = \sum_{x \leq a} p(x)$$

처럼 정의되고, F는 계단함수처럼 나타난다.



반면 random variable이 continuous라면, 아래 박스가 성립한다.

X가 continuous random variable이라면, nonnegative function인 **probability density function** $f(x)$가 존재하여

$$P\{X \in B\} = \int_B f(x)dx$$

이다. 이때, 모든 실수 a에 대하여

$$P\{X = a\} = \int_a^a f(x)dx = 0$$

이다.

또 전체 확률은 항상 1이어야 하기에, 아래 등식 또한 성립한다.

$$\int_{-\infty}^{\infty} f(x)dx = P\{X \in (-\infty, \infty)\} = 1$$

또 정의에 의하여 cumulative distribution function은

$$F(a) = \int_{-\infty}^{a} f(x)dx$$

처럼 표현가능하고,

$$P\{a \le X \le b\} = \int_a^b f(x)dx = F(b) - F(a)$$

이다. 우리는 여기에서 FTC를 이용하면

> cumulative distribution function F와 probability density function f에 대하여, 이들이 continuous 라면,
>
> $$\frac{d}{dx}F(x) = f(x)$$
>
> 이다.

를 얻는다.

8.6 Expectation and Variance

random variable X가 평균적으로 어떤 값을 가진다고 예측하고 싶을 때, 우리는 **expectation**(**expected value**)를 사용해줄 수 있다.

> random variable X가 discrete일 때에는 expectation $E[X]$가
>
> $$E[X] = \sum_{i=1}^{\infty} x_i P(X = x_i)$$
>
> 로 정의되며, continuous일 때는
>
> $$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$
>
> 으로 정의된다.

단, 확률변수 X가 항상 expectation을 가지는 것은 아니다. continuous random variable X가 아래의 probability density function을 가진다고 하자.

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

일단 이 함수가 probability density function으로써 잘 기능함은 독자의 몫에 맡긴다. ($-\infty$와 ∞에서 적분해 적분값이 1임과, 저 함수가 항상 nonnegative임을 보여주면 되겠다.)

하지만

$$E[X] = \int_{-\infty}^{\infty} \frac{1}{\pi}\frac{x}{1+x^2}dx = \left[\frac{1}{2\pi}\ln(1+x^2)\right]_{-\infty}^{\infty}$$

이기에, X의 expectation은 존재하지 않는다. 즉, 제시된 저 improper integral은 convergent가 아니다. 이는 discrete random variable에서도 마찬가지인데, 이는 8단원 내용이 아니므로 여기서 자세한 예시는 들지 않으려 한다. 어찌하였든 random variable이 expectation을 가지지 않을 수 있음을 기억하자.

또한 expectation에 대해 아래 박스도 성립한다.

> $$E[g(X)] = \begin{cases} \sum_{i=1}^{\infty} g(x_i)P(X = x_i) & (X\text{: discrete}) \\ \int_{-\infty}^{\infty} g(x)f(x)dx & (X\text{: continuous}) \end{cases}$$
>
> $$E[aX + bY] = aE[X] + bE[Y]$$

이 expectation $E[X] = \mu$는 $E[(X-c)^2]$을 최소화하는 c의 값이기도 하다. 즉, expectation은 제곱오차 $(X-c)^2$의 평균을 최소화한다는 것이다. 아래의 식으로부터 이를 얻을 수 있다.

$$\begin{aligned} E[(X-c)^2] &= E[(x-\mu+\mu-c)^2] \\ &= E[(X-\mu)^2 + 2(\mu-c)(X-\mu) + (\mu-c)^2] \\ &= E[(X-\mu)^2] + 2(\mu-c)(E[X]-\mu) + (\mu-c)^2 \\ &\geq E[(X-\mu)^2] \end{aligned}$$

이러한 성질을 감안하면, expectation을 알 때 $E[(X-\mu)^2]$의 값이 궁금해진다.

X가 expectation μ를 가지는 random variable일 때, X의 variance는 아래와 같다.

$$Var(X) = E[(X-\mu)^2]$$

만약 여기서 더 나아간다면 우리는

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - (E[X])^2 \end{aligned}$$

을 얻을 수도 있다. 중학교 재학 중 들어보았던 제평평제가 바로 이러한 과정을 거쳐 등장한 것이다. 주의할 것은 variance는 expectation이 있는 random variable에만 적용된다는 것이다. 전 페이지에서 살펴보았듯 expectation이 없는 random variable에 대해서는 variance조차 정의되지 않는다. 또 variance는 항상 음이 아니라는 사실을 기억하자.

한편 expectation에 대해 $E[aX+b] = aE[X]+b$가 성립했었으므로, variance에 대해서는 아래가 성립한다.

$$Var(aX+b) = a^2 Var(X)$$

상수에 대해서는 $a = 0$이므로, $Var(b) = 0$임을 확인해줄 수 있다. 다르게 말하면 variance가 0이라는 것은 곧 해당 random variable이 확률 1로 특정 값을 가진다는 것이다. 한편

random variable X의 **standard deviation**은 아래와 같다.

$$sd(X) = \sqrt{Var(X)}$$

가끔씩은 median이라는 개념을 사용하기도 한다.

random variable의 median은

$$F(a) = 0.5$$

인 a값을 의미한다.

median은 $E[|X-c|]$를 최소화하는 c의 값이기도 하다. 즉 median은 절댓값오차 $|X-c|$의 평균을 최소화한다.

이외의 다른 개념, 예를 들어 independence나 각종 distribution은 해마다, 선생님들마다 가르치시는지의 여부가 다르다. 따라서 여기서는 생략하고, 연습문제에서 살짝씩 맛보고자 한다.

8.7 연습문제

문제 8. 1. *R은 $y = 0$과 $y = \sqrt{1-x^2}$에 의해 둘러싸인 영역이다. R을 $y = x - 1$로 돌렸을 때 얻어지는 입체의 부피를 구하시오.*

문제 8. 2. *$y = \sqrt{1-x^2}$를 $0 \le x \le \frac{1}{2}$에서 x축으로 회전시킬 때, 회전체의 표면적을 구하시오.*

문제 8. 3. $\bar{x}$를 $a \le x \le b$에서 $y = f(x) \ge 0$의 그래프 아래에 있는 영역의 *centroid*가 갖는 x좌표라 하자. 이때

$$\int_a^b (cx + d) f(x) dx = (c\bar{x} + d) \int_a^b f(x) dx$$

임을 보여라.

문제 8. 4. 두 그래프 $y = x^3 - x$와 $y = 1 - x^4$로 둘러싸인 영역의 *centroid*를 구하여라.

문제 8. 5. $\mathcal{R}$은 $-1 \le x \le 1$과 $2 \le y \le 4$에 의해 둘러싸인 직사각형 영역이다. 이를 $3x-4y=0$을 기준으로 회전시켰을 때 얻어지는 입체의 부피를 구하여라.

문제 8. 6. x축보다는 위에 있고, $y = 1 - x^n$보다는 아래에 있는 영역의 *centroid*를 구하여라. 이때, n은 양의 짝수라고 하자. 만약 $n \to \infty$라면, *centroid*는 어디로 수렴하는가?

문제 8. 7. $y = \frac{1}{2}x$를 축으로 하여 $0 \leq y \leq \sqrt{5 - x^2}$, $0 \leq x \leq 1$로 표현되는 영역을 회전시킬 때, 표면적을 구하시오.

문제 8. 8. *한 변의 길이가 10인 정사각형을 잘라, 오른쪽 아래 꼭짓점만을 포함하는 넓이가 30인 직각 삼각형과 나머지 부분으로 나누었다. 이때 나머지 부분의 centroid는 나머지 부분에 온전하게 남아있는 두 변 중 왼쪽 변에 4만큼의 거리로 떨어져 있다고 한다. 그렇다면 centroid는 위쪽 변과는 얼마만큼의 거리를 가지겠는가?*

문제 8. 9. $y = \frac{1}{x^2+1}$, $y = 0, x = 0, x = 1$로 *둘러싸인 영역의 centroid를 구하여라.*

문제 8. 10. $y^2 = x^3 - x^4$으로 *주어지는 곡선을 생각하자. 이 곡선에 의해 생성되는 닫힌 영역의 centroid를 구하여라.*

문제 8. 11. *$y^2 = x^3 - x^4$으로 주어지는 곡선을 생각하자. 이 곡선에 의해 생성되는 닫힌 영역을 $(0, 1)$을 지나는 어떤 직선 l에 대해 회전시켰다. l이 양의 기울기를 가진다고 할 때, 회전체의 최대 부피를 구하여라.*

문제 8. 12. *$f(t)$가 $[0,1]$에서의 differentiable function이고 $f(0) = 0, f(1) = 1$이다. 또한 $f'(t) \geq 0$이 $t \in (0,1)$에서 성립한다. 아래 부등식이 성립함을 보여라.*

$$\sqrt{2} \leq \int_0^1 \sqrt{1 + [f'(t)]^2} dt \leq 2$$

문제 8. 13. *X가 continuous random variable일 때,*

$$f(x) = \begin{cases} xe^{-x} & x \geq 0 \\ 0 & otherwise \end{cases}$$

라고 한다. $f(x)$가 probability density function이 될 수 있음을 보이고, 그 cumulative distribution function 을 구하여라.

문제 8. 14. *X가 continuous random variable일 때, 그 probability density function이다.*

$$f(x) = \begin{cases} xe^{-x} & x \geq 0 \\ 0 & otherwise \end{cases}$$

라고 한다. $Y = X^2$으로 새로운 random variable Y를 정의할 때 Y의 probability density function을 구하고, X, Y의 expectation과 variance를 구하여라.

문제 8. 15. *영역*

$$S = \{(x, y) | x \geq 0, y \leq x^2 + y^2 \leq 4y\}$$

의 centroid를 구하여라.

문제 8. 16. *random variable X가 아래의 probability density function을 가진다.*

$$f(x) = \begin{cases} \frac{1}{2} & 0 \leq x \leq 1 \\ -\frac{1}{4}x + k & 1 < x \leq 3 \\ 0 & otherwise \end{cases}$$

이때 k, $E[X], Var(X)$를 구하시오.

문제 8. 17. *곡선* $y = \frac{x^4}{16} + \frac{1}{2x^2}$*을* $1 \leq x \leq 2$*의 범위에서* x*축을 기준으로 회전시켰을 때, 회전체의 표면적을 구하여라.*

문제 8. 18. *median은 continuous random variable* X*의* $E[|X - c|]$*를 최소화하는* c*의 값임을 보여라. 단,* X*의 cumulative distribution function* F*와 probability density function* f*는 continuous이다.*

문제 8. 19. *양의 정수* n*에 대하여* $y = x^n$*과* $y = 0, x = 1$*으로 둘러싸인 영역의 centroid를* $(\overline{x_n}, \overline{y_n})$*라고 하자.*

$$\lim_{n\to\infty} \overline{y_n} - \overline{x_n}$$

을 구하여라.

문제 8. 20. *random variable* X*는 오직 1, 2, 3만을 값으로 가진다고 한다. 만약* $E[X] = 2$*라면,* $Var(X)$*를 최소화시키는* $p_i = P\{X = i\}$*의 값은 각각 어떻게 될 것인가?* $Var(X)$*를 최대화시키려면 어떻게 해야 할까?*

문제 8. 21. *인기 예능 신서유기에서는* n*명의 멤버를 데리고 90년대 노래 퀴즈를 진행한다. 각각이 정답을 맞출 확률은* p*이며, 각자의 결과에 독립적이다. 나PD는 최소 반 이상이 정답을 맞췄을 때 식사를 제공하기로 했다. 이때 식사를 제공받으면* $X_p = 1$*, 제공받지 못하면* $X_p = 0$*이라 하자. 즉* X_p*는 식사 제공에 대한 indicator variable이다.*

1) p*가 어느 값 이상이어야 5명의 멤버가 도전했을 때 3명의 멤버가 도전했을 때보다 더 성공 확률이 높을 것인가?*

2) p*가 어느 값 이상이어야* $2k+1$*명의 멤버가 도전했을 때* $2k-1$*명의 멤버가 도전했을 때보다* $E[X_p]$*가 더 높을 것인가?*

문제 8. 22. *나PD는 다음 게임으로* n*명의 멤버들을 방에 불러 두고 그들의 모자를 모두 모아 섞었다. 그 다음, 연장자 순으로 무작위의 모자를 고르고 가져갔다.* X*를 자신의 모자를 되찾은 멤버의 수라고 하자.* $E[X] = 1$*임을 보여라.*

문제 8. 23. *배구선수 김연경의 리시브효율은* p*이며, 연달아서 서브를 계속 받고 있다. 또한, 각 서브리시브의 성공확률인* p*는 이전/이후 서브 시도의 영향을 받지 않는다. random variable* X*를 성공까지 필요한 시도 횟수라고 하자. 즉,* X*가* k*라는 말은* $k-1$*번의 리시브 실패 후* k*번째 서브를 잘 받아냈다는 것을 의미한다.*

X*의 expectation은 얼마인가?*

문제 8. 24. *평소 연습량이 많은 김연경 선수는* r*번의 성공을 얻기 전까지 리시브 훈련을 계속한다. 이처럼, 특정 성공 횟수를 얻기 위한 시행의 수를* Y*라고 할 때,* Y*는* ***negative geometric distribution****을 따른다고 한다.*

1) $k = r, r+1, \cdots$ *에 대하여* $P(Y = k)$*을 구하여라.*

2) $E[Y] = r/p$*임을 보여라.*

문제 8. 25. *continuous random variable* X*에 대하여,* $E[X]$*가 존재한다고 가정하자.*

$$E[X] = \int_0^\infty P(X > x)dx - \int_0^\infty P(X < -x)dx$$

임을 보여라. 단, $E(X)$*가 존재하면* $\lim_{t\to\infty} tF(-t) = 0$ *임은 증명 없이 사용해도 괜찮다.*

문제 8. 26. X*가 continuous random variable이며,* $a \le X \le b$*이고* $E[X] = \mu$*라고 하자.*

(1) $a \le \mu \le b$*임을 보여라.*

(2) $V(X) \le \frac{1}{4}(b-a)^2$ *임을 보여라.*

문제 8. 27. *소방서가 수직선 위에 있다. 불의 위치가 $f(x)$라는 probability density function을 가진 random variable에 따라 분포해 있으며, $\int_{-\infty}^{\infty}|x|f(x)dx < \infty$를 만족한다고 하자. 그렇다면, 불로부터 소방서 사이의 거리의 expectation을 최소화할 수 있는 소방서의 위치는 어디인가?*

문제 8. 28. *random variable R은*

$$f_R(r) = \frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}$$

이라는 probability density function을 가질 때 Rayleigh distribution이라고 부른다. $E[R]$을 σ를 이용해 표현하여라.

단, $\int_{-\infty}^{\infty} e^{-x^2/2}dx = \sqrt{\frac{\pi}{2}}$임을 이용하여도 좋다.

문제 8. 29. $X_1, X_2, \cdots, X_n$은 독립인 *random variable이며, 그들의 probability density function은 모두*

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & otherwise \end{cases}$$

라고 하자. $M = max(X_1, X_2, \cdots, X_n)$*이라고 정의할 때,* M*의 probability density function를 구하여라. 이때 random variable끼리 독립이라는 것은, 각 random variable이 특정한 값을 가질 확률이 다른 random variable이 어떤 값을 가지는지에 무관하다는 것이다.*

문제 8. 30. $X_1, X_2, \cdots$는 독립이며 동일한 *probability density function를 가지는 continuous random variable의 수열이다.* $N \ge 2$를

$$X_1 \ge X_2 \ge \cdots \ge X_{N-1} < X_N$$

인 N*이라 정의하자. 즉, 감소를 멈추는 첫 점이라고 생각하자. 이때 random variable끼리 독립이라는 것은, 각 random variable이 특정한 값을 가질 확률이 다른 random variable이 어떤 값을 가지는지에 무관하다는 것이다.*

1) $P(N \ge n)$*의 값을 구하여라.*

2) $e = 1 + \dfrac{1}{2!} + \dfrac{1}{3!} + \cdots$ *임을 알 때,* $E[N] = e$*임을 보여라.*