확률 및 통계학

- 연속형 확률분포(6.1~6.4) -

송 영 준(youngjun@pel.smuc.ac.kr)

상명대학교 프로토콜공학연구실

목 차

- 연속형 균일분포
- 정규분포
- 표준정규분포

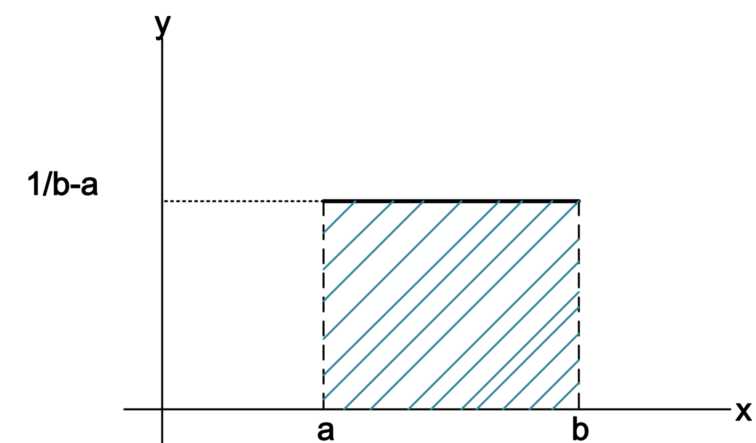
연속형 균일분포

- 균일분포
  - 과거의 경험이 미래를 예측하는 어떠한 영향도 미치지 않으며, 나타날 가능성이 모두 동일한 분포
    - 이산형 균일분포, 연속형 균일분포

    - 이산형 균일분포
      - 확률 변수 X가 $x_1, x_2, \ldots, x_k$의 각 값을 취할 확률이 동일한 경우
    - 연속형 균일분포
      - 확률 변수 X가 구간$[A, B]$에서 각 값을 취할 확률이 동일한 경우

연속형 균일분포

- 정의
  - 구간[A,B]에서 정의되는 연속형 균일확률변수 X의 밀도함수
    - $f(x;A,B)=\begin{cases}\frac{1}{B-A}, A\le x\le B\\ 0, \text{다른 곳에서}\end{cases}$
    - 밀도함수는 밑변이 b-a이고, 높이는 $\frac{1}{b-a}$로 일정한 직사각형
      - 전체 확률(면적)은 1
      - 동일한 확률이 분포하여 일자형 그래프가 나타남



연속형 균일분포

- 예제6.1
  - 어느 회사의 대형 회의실은 4시간을 초과하여 사용할 수 없고, 그 회의실에서는 긴 회의와 짧은 회의가 자주 열리며, 회의시간X는 구간[0,4]에서 정의되는 균일분포로 가정
    - (a)밀도함수
      - $f(x) = \begin{cases} \frac{1}{4}, 0 \le x \le 4 \\ 0\text{, 다른 곳에서} \end{cases}$
    - (b)어떤 회의가 최소한 3시간 이상 계속될 확률
      - $P[X \ge 3] = \int_3^4 f(x)dx = \frac{1}{4}$

연속형 균일분포

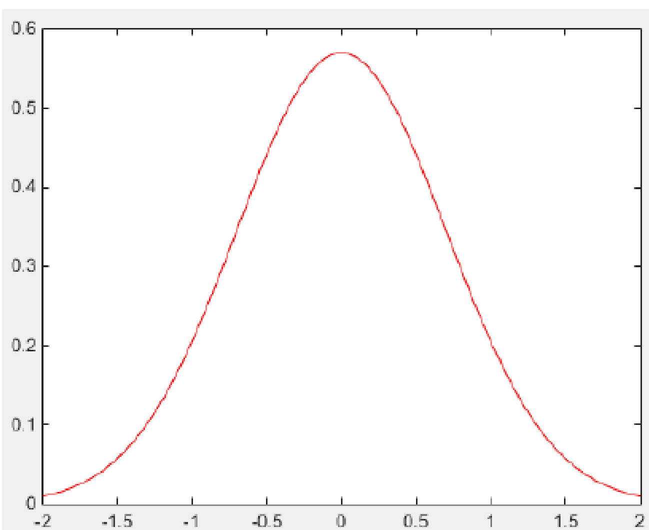
- 평균, 분산
  - $\mu = \frac{A+B}{2}, \sigma^2 = \frac{(B-A)^2}{12}$
- 증명
  - $\mu = \int_a^b xf(x)dx = \int_a^b x\frac{1}{b-a} = \frac{1}{b-a}\int_a^b xdx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$
  - $\sigma^2 = \int_a^b x^2 f(x)dx - m^2 = \int_a^b x^2\frac{1}{b-a}dx - \left(\frac{b+a}{2}\right)^2 = \frac{x^3}{3(b-a)}\Big|_a^b - \frac{(b+a)^2}{4} = \frac{(b-a)\left(b^2+ba+a^2\right)}{3(b-a)} - \frac{a^2+2ab+b^2}{4} = \frac{\left(b^2+ba+a^2\right)}{3} - \frac{a^2+2ab+b^2}{4} = \frac{(b-a)^2}{12}$

정규분포

- 정규분포(Normal Distribution)
  - 연속확률분포의 한 종류로 통계학에서 가장 많이 활용되는 확률분포
  - 확률밀도함수가 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$를 따르는 확률분포
    - $-\infty < \mathrm{x} < \infty$
  - 확률밀도함수가 종 모양을 띰
  - 정규확률변수의 확률분포는 평균μ와 표준편차σ에 의해서 결정되기 때문에 정규확률변수X의 밀도함수를 $n(x;\mu,\sigma)$로 표시
  - 자연과학, 기업, 각종 연구분야에서 발생하는 여러 현상들을 근사적으로 기술하는데 사용

정규분포

- 정규분포(Normal Distribution)
  - 정의
    - 확률변수 X가 음의 무한대에서 양의 무한대사이의 어떠한 값도 취할수 있는 연속확률분포로서 μ(기대값)와 σ(표준편차) 2개의 모수에 의해 특징지어지는 분포
      - $n(x;\mu,\sigma) = f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, -\infty < x < \infty$
      - $\pi = 3.14159\ldots, e = 2.71828\ldots$



정규분포

- 정규분포(Normal Distribution)
  - 특징
    - $x = \mu$에서 곡선이 최대값이 되며, 또한 최빈값(mode)을 가짐
    - 곡선은 평균μ를 지나는 수직축에 대하여 대칭
    - 곡선은 $x = \mu \pm \sigma$에서 변곡점이 나타남
    - 평균에서 멀어질수록, 정규곡선은 수평축에 접근
    - 곡선과 수평축사이의 총 면적은 1(확률)

표준정규분포

- 표준정규분포(Standard Normal Distribution)
  - 정규분포의 평균과 표준편차에 따라 모양과 위치가 다르기 때문에 두 분포의 성격을 비교하거나, 특정 정규분포에서 확률을 계산하기 위해서 정규분포의 평균과 표준편차를 표준화하여 표준적인 정규분포를 만듦
  - 정규밀도함수의 적분의 어려움을 피하고, 쉽게 이용할 수 있도록 여러 구간에 대한 면적을 구할 수 있는 표가 필요
    - 표준정규분포표

표준정규분포

- 표준정규분포(Standard Normal Distribution)
  - 표준정규분포표

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5793
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.6365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319

표준정규분포

- 표준정규분포(Standard Normal Distribution)
  - 정의
    - 평균이 0이고, 분산이 1인 정규확률변수의 분포를 표준정규분포라 함
    - 모든 정규확률변수 X를 표준정규확률변수 Z로 변환 가능
      - $Z = \frac{X-\mu}{\sigma}$

표준정규분포

- 예제6.2
  - 표준정규분포가 주어졌을 때, 면적을 구하라
    - (a)$z = 1.84$의 오른쪽 면적
      - $z = 1.84 = 0.9671\,, 1 - 0.9671 = 0.0329$
    - (b)$\mathrm{z} = -1.97$과 $z = 0.86$사이의 면적
      - $z = 0.86 = 0.8051, z = -1.97 = 0.0244$
      - $0.8051 - 0.0244 = 0.7807$

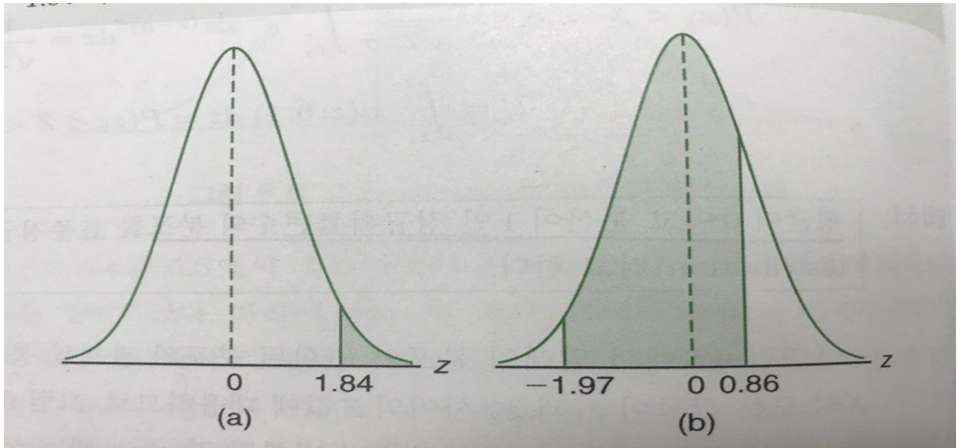


그림 6.9 예제 6.2에 대한 면적

표준정규분포

- 예제6.3
  - 표준정규분포가 주어졌을 때, 다음 각 경우에 대한 k값
    - (a)$P(z > k) = 0.3015$
      - 1-0.3015=0.6985, 0.6985=0.52
    - (b)$P(k < Z < -0.18) = 0.4197$
      - z=-0.18=0.4286, 0.4286-0.4197=0.0089
      - 0.0089=-2.37

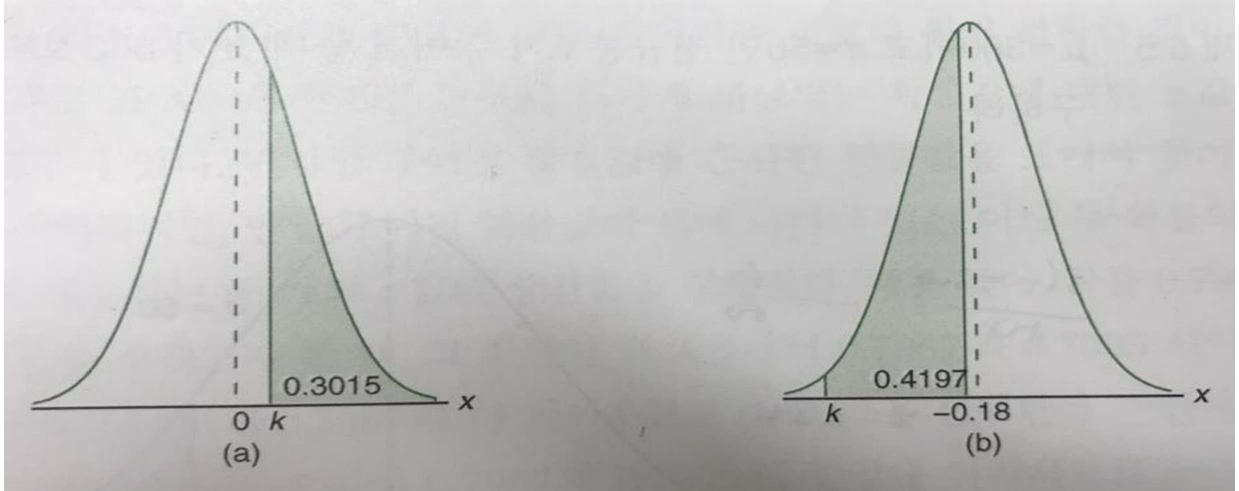


그림 6.10 예제 6.3에 대한 면적

표준정규분포

- 예제6.4
  - $\mu = 50, \sigma = 10$인 정규분포가 주어졌을 때, X가 45와 62 사이의 값을 취할 확률
    - $z_1 = \frac{X-\mu}{\sigma} = \frac{45-50}{10} = -0.5, \quad z_2 = \frac{X-\mu}{\sigma} = \frac{62-50}{10} = 1.2$
    - $P(-0.5 < Z < 1.2) = P(Z < 1.2) - P(Z < -0.5) = 0.8849 - 0.3085 = 0.5764$

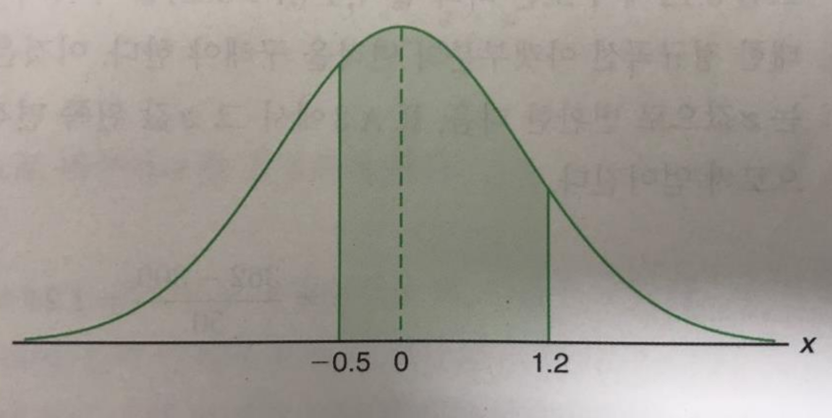


그림 6.11 예제 6.4에 대한 면적

표준정규분포

- 예제6.5
  - $\mu = 300, \sigma = 50$인 정규분포가 주어졌을 때, X가 362보다 큰 값을 취할 확률
    - $P(X > 362)$
    - $Z = \frac{362-300}{50} = 1.24, \ \ P(Z > 1.24)$
    - $Z = 1.24 = 0.8925, \ 1 - 0.8925 = 0.1075$

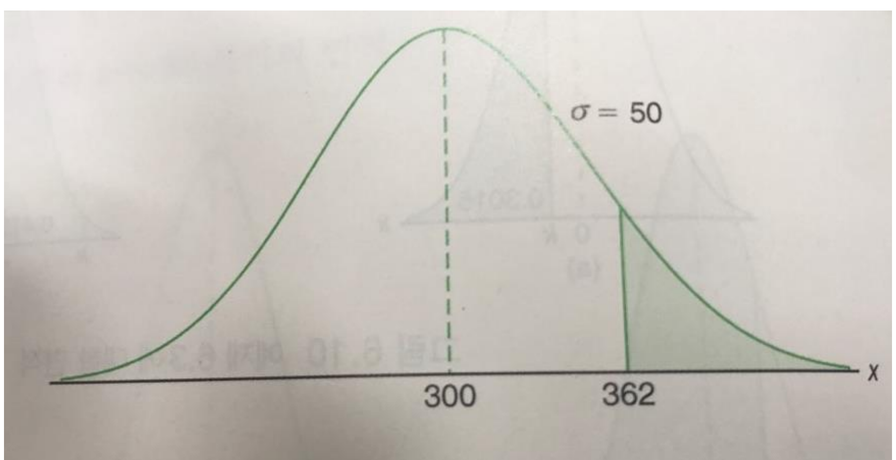


그림 6.12 예제 6.5에 대한 면적

표준정규분포

- 예제6.6
  - $\mu = 40, \sigma = 6$인 정규분포가 주어졌을 때, 다음을 구하라
    - (a)왼쪽 면적이 전체 면적의 45%가 되는 x
      - $P(Z < t) = 0.45, t = -0.13$
      - $\therefore x = \sigma z + \mu = (6)(-0.13) + 40 = 39.22$
    - (b)오른쪽 면적이 전체 면적의 14%가 되는 x
      - $1 - 0.14 = 0.86, P(Z < t) = 0.86, t = 1.08$
      - $\therefore x = \sigma z + \mu = (6)(1.08) + 40 = 46.48$

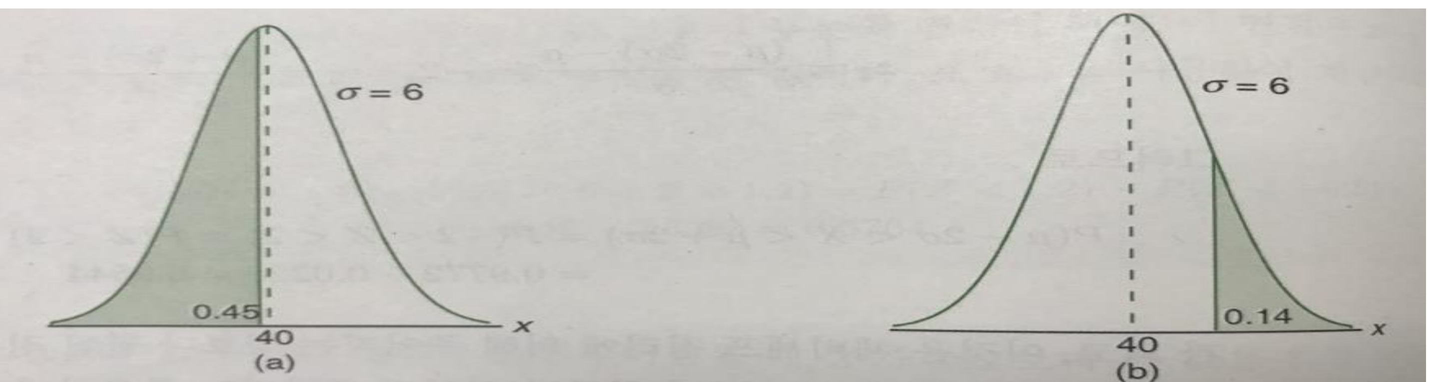
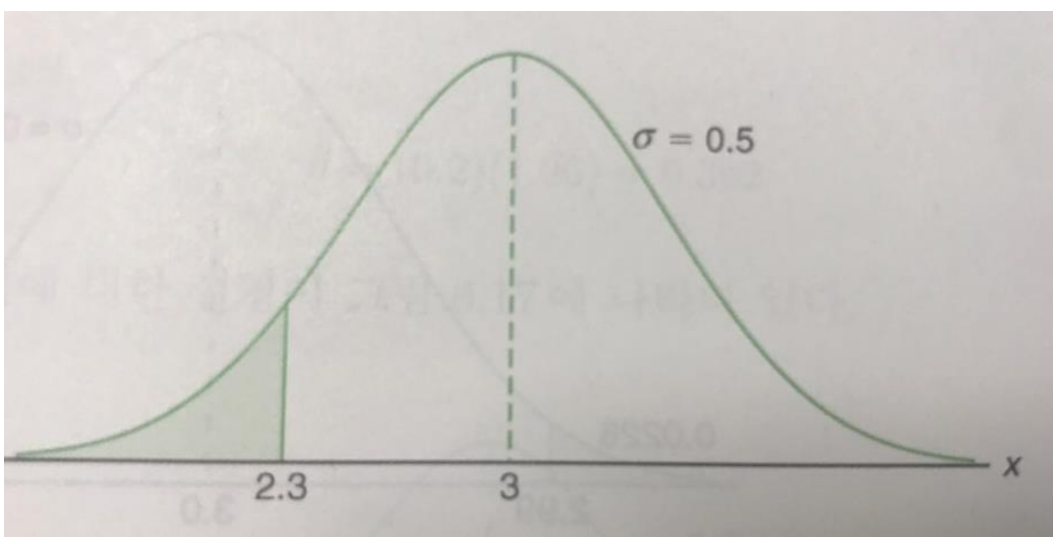


그림 6.13 예제 6.6에 대한 면적

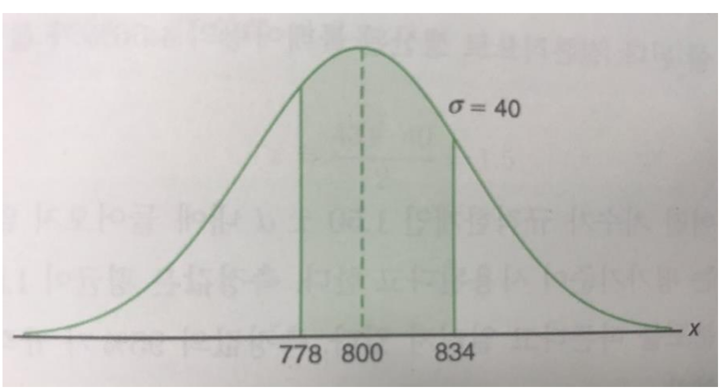
표준정규분포

- 예제6.7
  - 어느 축전지의 평균수명이 3년이고 표준편차가 0.5이고, 축전지의 수명이 정규분포를 따를 때, 임의로 주어진 전지의 수명이 2.3년보다 짧을 확률
    - $P(X < 2.3)$
    - $Z = \frac{X-\mu}{\sigma} = \frac{2.3-3}{0.5} = -1.4,\ P(Z < -1.4) = 0.0808$



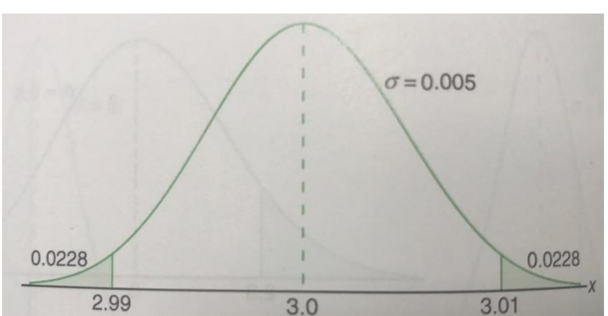
표준정규분포

- 예제6.8
  - 어느 전기회사에서 평균수명이 800시간이고 표준편차가 40시간인 정규분포의 수명분포를 가지는 전구를 생산할 때, 임의로 선정된 전구의 수명이 778시간과 834시간 사이에 있을 확률
    - $P(778 < X < 834)$
    - $Z_1 = \frac{X-\mu}{\sigma} = \frac{778-800}{40} = -0.55, \quad Z_2 = \frac{X-\mu}{\sigma} = \frac{834-800}{40} = 0.85$
    - $P(-0.55 < Z < 0.85) = P(Z < 0.85) - P(Z < -0.55) = 0.8023 - 0.2912 = 0.5111$



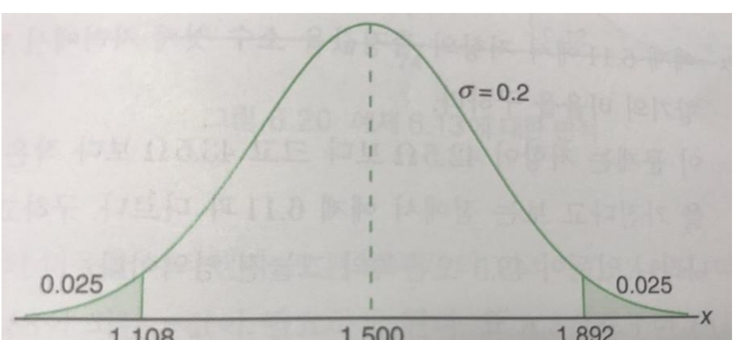
표준정규분포

- 예제6.9
  - 어느 볼베어링의 직경에 대한 규격한계를 3.0±0.01Cm로 정하였고, 이 규격한계를 벗어나는 부품은 불합격으로 처리할 때, 볼베어링의 직경은 평균이 3.0, 표준편차 0.005인 정규분포를 따른다면 생산된 제품 중 불합격으로 처리되는 것은얼마나 되겠는가?
    - $Z_1 = \frac{X-\mu}{\sigma} = \frac{2.99-3.0}{0.005} = -2.0, \qquad Z_2 = \frac{X-\mu}{\sigma} = \frac{3.01-3.0}{0.005} = 2.0$
    - $P(Z < -2.0) = 0.0228,\ 2(0.0228) = 0.0456$
    - 생산된 볼베어링의 4.56%가 불합격품으로 처리



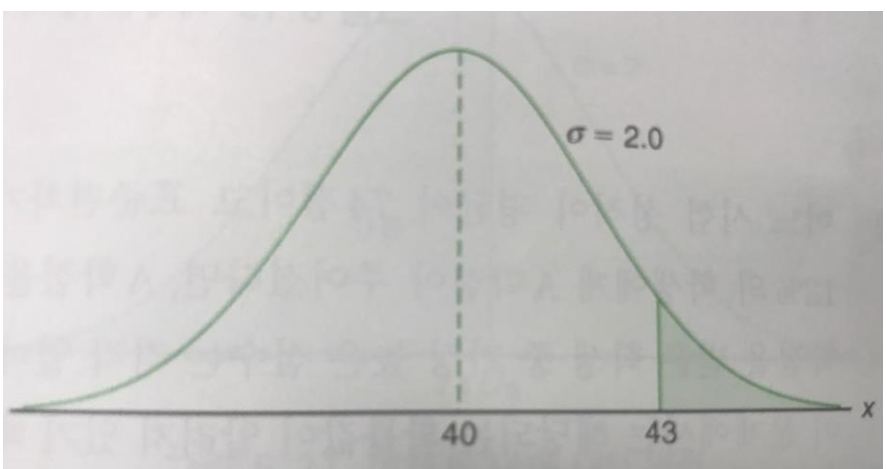
표준정규분포

- 예제6.10
  - 어떤 치수가 규격한계인 1.50$\pm d$내에 들어오지 않으면 부품을 불합격시키는 평가기준을 사용할 때, 측정값은 평균이 1.50이고 표준편차가 0.2인 정규분포에서 측정값의 95%가 규격한계 내에 들도록 d값을 결정하라
    - 합격 : 0.95
    - $P(X > d) = 0.5 - 0.0475 = 0.025, \quad P(-1.96 < Z < 1.96) = 0.95$
    - $\therefore d = \sigma z + \mu = (0.2)(1.96) + 1.50 = 0.392$



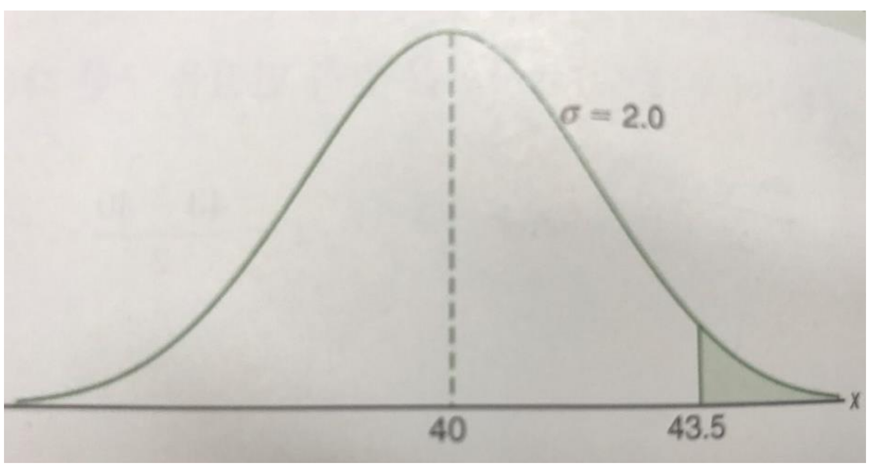
표준정규분포

- 예제6.11
  - 평균저항이 40Ω이고 표준편차가 2Ω인 저항기를 만드는 기계가 있을 때, 저항이 정규분포를 따르고, 43Ω이 넘는 저항을 가지게 되는 저항기는 몇 퍼센트나 되겠는가?
    - $P(X > 43)$
    - $P\left(Z > \frac{43-40}{2}\right) = P(Z > 1.5) = 1 - P(Z < 1.5) = 1 - 0.9332 = 0.0668$



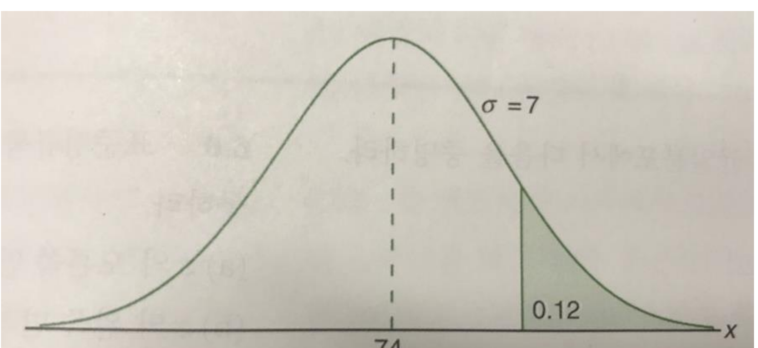
표준정규분포

- 예제6.12
  - 예제 6.11에서 저항의 측정값을 소수 첫째 자리에서 반올림할 때 43Ω이 넘는 저항기의 비율을 구하라
    - $P(X > 43.5)$
    - $P\left(Z > \frac{43.5-40}{2}\right) = P(Z > 1.75) = 1 - P(Z < 1.75) = 1 - 0.9599 = 0.0401$



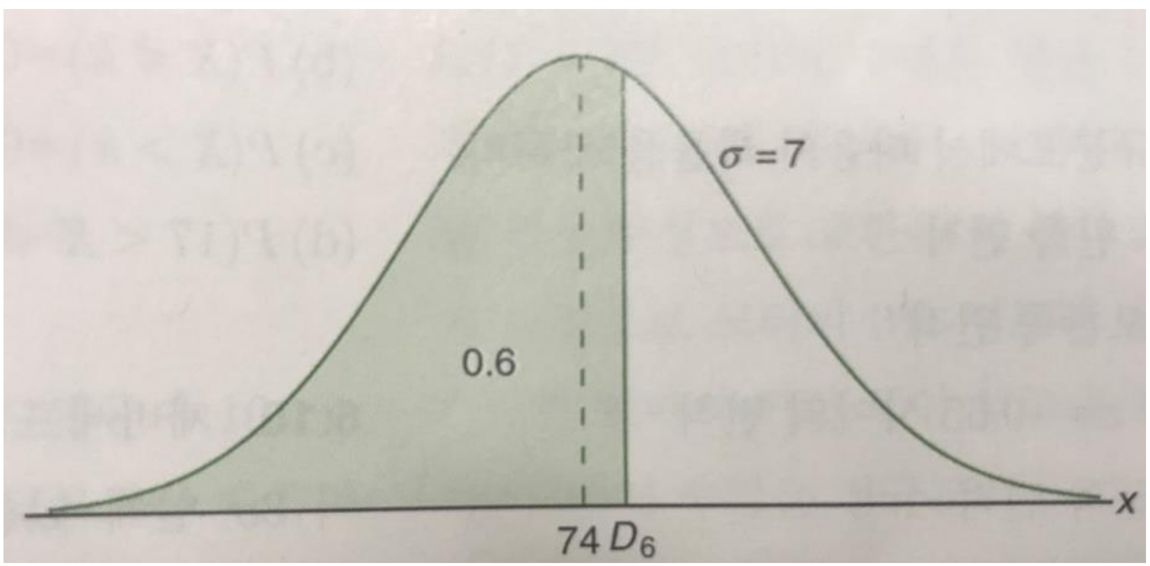
표준정규분포

- 예제6.13
  - 어느 시험 성적이 평균이 74점이고 표준편차가 7인 정규분포를 따른다고 할 때, 12%의 학생에게 A학점이 주어졌다면, A학점을 받은 학생 중 가장 낮은 점수와 B학점을 받은 학생 중 가장 높은 점수는 각각 얼마나 되겠는가?
    - 학점A는 확률밀도함수 오른쪽 부분에서 0.12부분을 나타내고, 왼쪽 면적이 0.88이 되는 z값을 구하면
      - $P(Z < 1.18) = 0.88$
      - $X = (7)(1.18) + 74 = 82.26$
      - A학점 중 가장 낮은 점수는 83점, B학점 중 가장 높은 학점은 82



표준정규분포

- 예제6.14
  - 예제 6.13에서 제 6십분위수를 구하라
    - 제 6십분위수를 D_6라 하면, 아래 그림과 같이 왼쪽 면적이 총 면정의 60%가 되는 x값이 된다. 표준정규분포표로부터 $P(Z < 0.25) \approx 0.6$이 되므로, $x = (7)(0.25) + 74 = 75.75$
      - D_6는 75.75, 성적의 60%가 75점 이하



감사합니다!