3. 정규분포

정규분포

정규분포가 통계학에서 중요시 되는 이유

① 자연현상의 대부분의 값은 거의 다 근사적으로 정규분포 하기 때문에

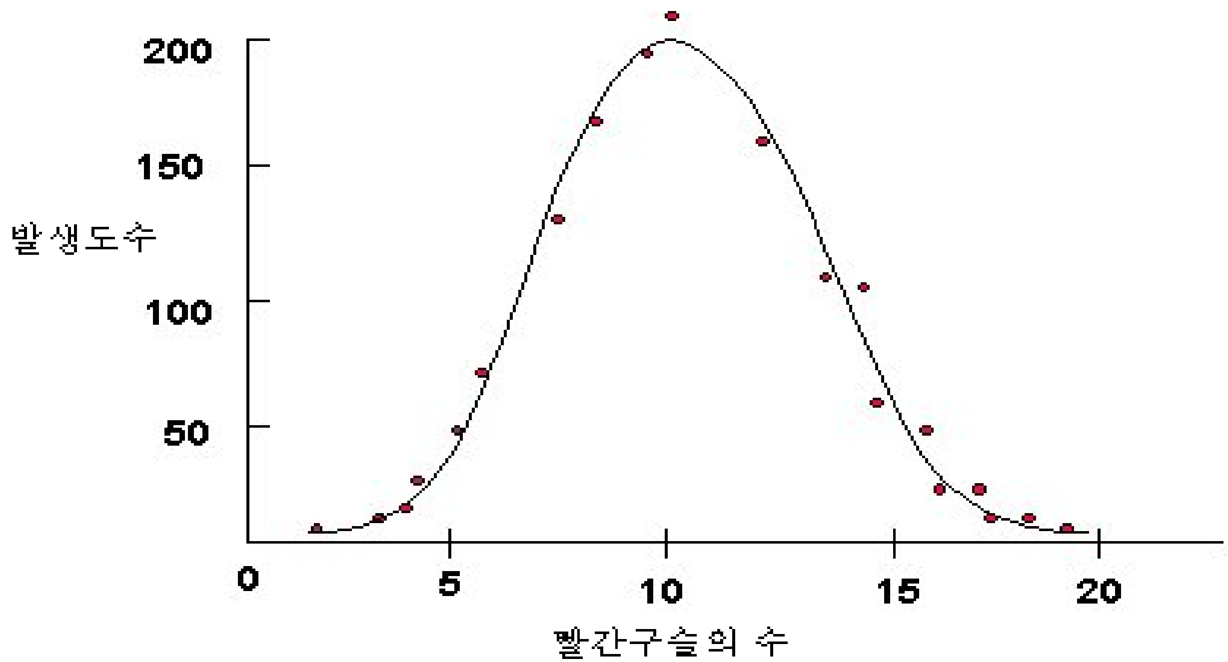
② 정규분포를 하지 않는 측정치들에 있어서도 측정척도에 대해 간단한 대수변환에 의해서 근사적으로 정규분포를 하도록 유도할 수 있기 때문에

③ 정규분포는 수학적으로 다루기 쉽기 때문에

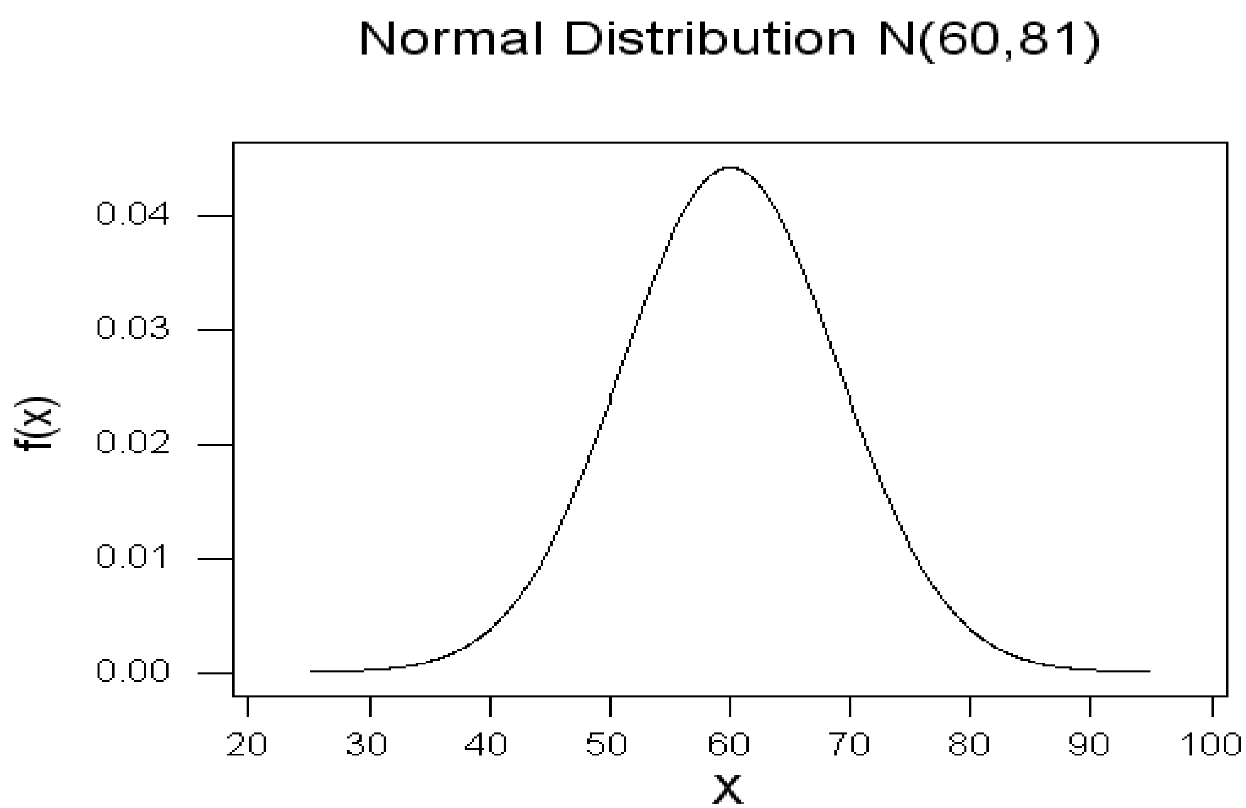
④ 모집단이 정규분포 하지 않아도 표본평균값 X는 표본의 크기가 클 때에는 근사적으로 정규분포 하기 때문에

정규분포 예(1)

- 2000회의 실험결과에 대한 도수분포를 아래와 같이 그려보면 아래와 같이 정규분포를 이룬다는 것을 알 수 있다.



정규분포 곡선


Normal Distribution N(60,81)
f(x)
0.04
0.03
0.02
0.01
0.00
20
30
40
50
60
70
80
90
100
X

정규분포

- 정규분포의 확률밀도함수

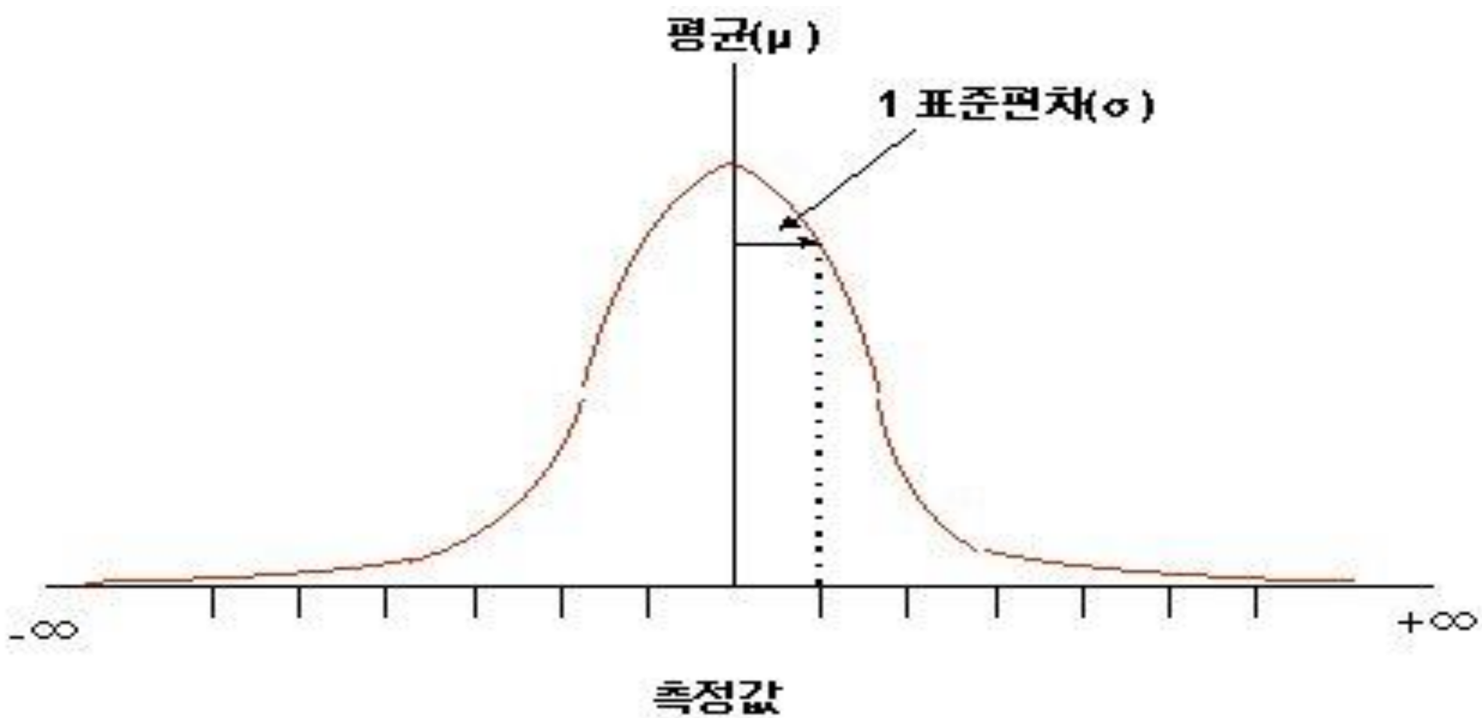
$$f(x;\mu,\ \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

단, $\pi = 3.141592\cdots\cdots$, $e = 2.7128\cdots\cdots$

- 정규분포는 연속확률분포로서 그 함수 f(χ)는 확률곡선의 높이를 나타내고, 그 높이는 밀도임
- 정규곡선의 모양은 평균(μ)과 분산(σ^2)에 의하여 결정됨(평균, 분산을 제외한 것은 모두 상수이므로)

정규분포의 특징

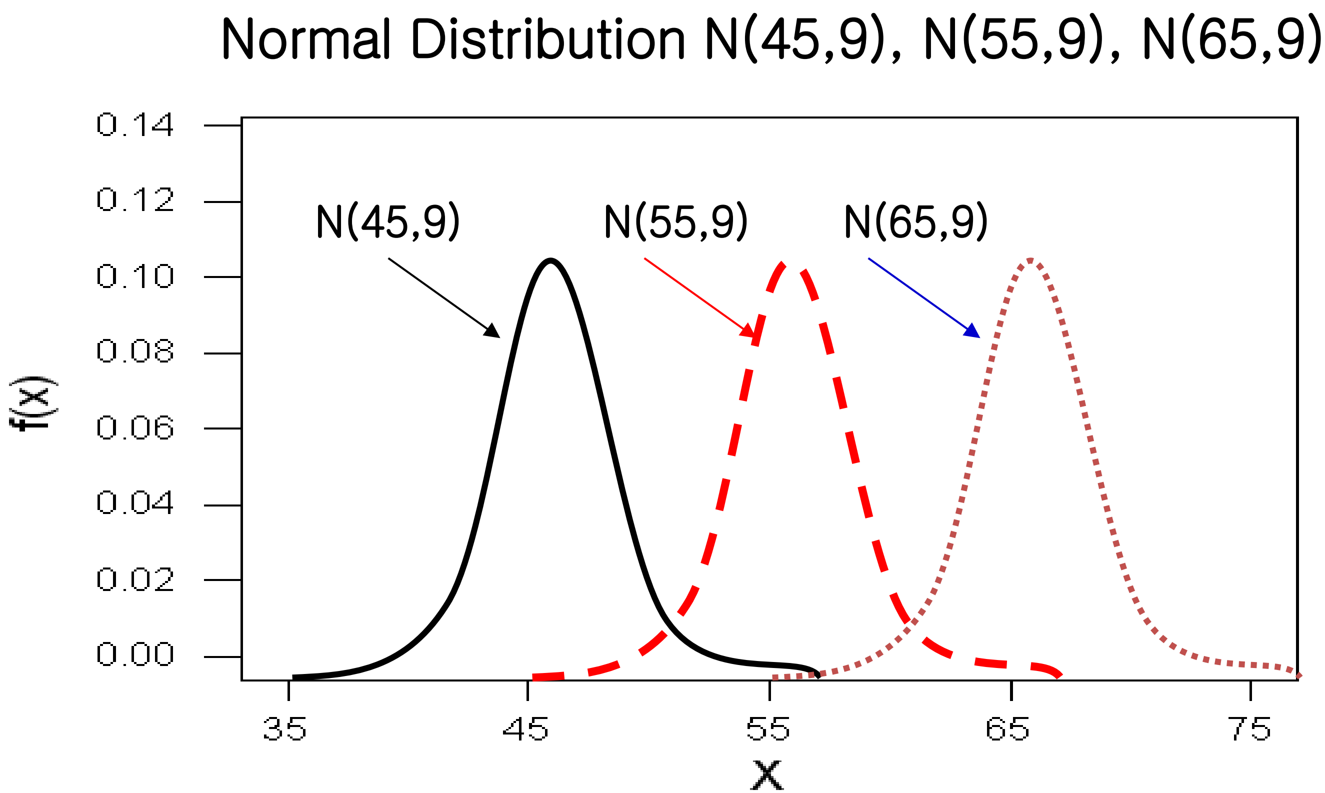
- 일반적인 정규분포 곡선과 μ와 σ의 관계



원광대학교
WONKWANG UNIVERSITY

정규분포의 특징(2)

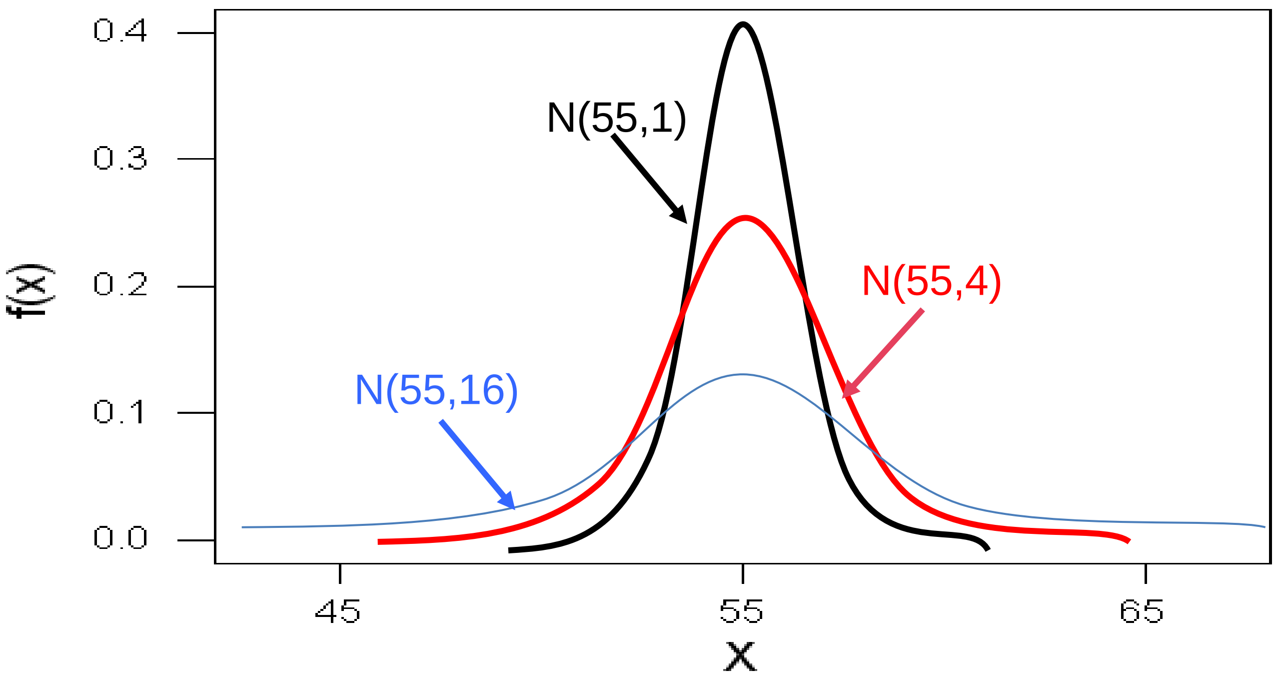
- 평균이 다르고 분산이 같은 분포의 형태



정규분포의 특징(2)

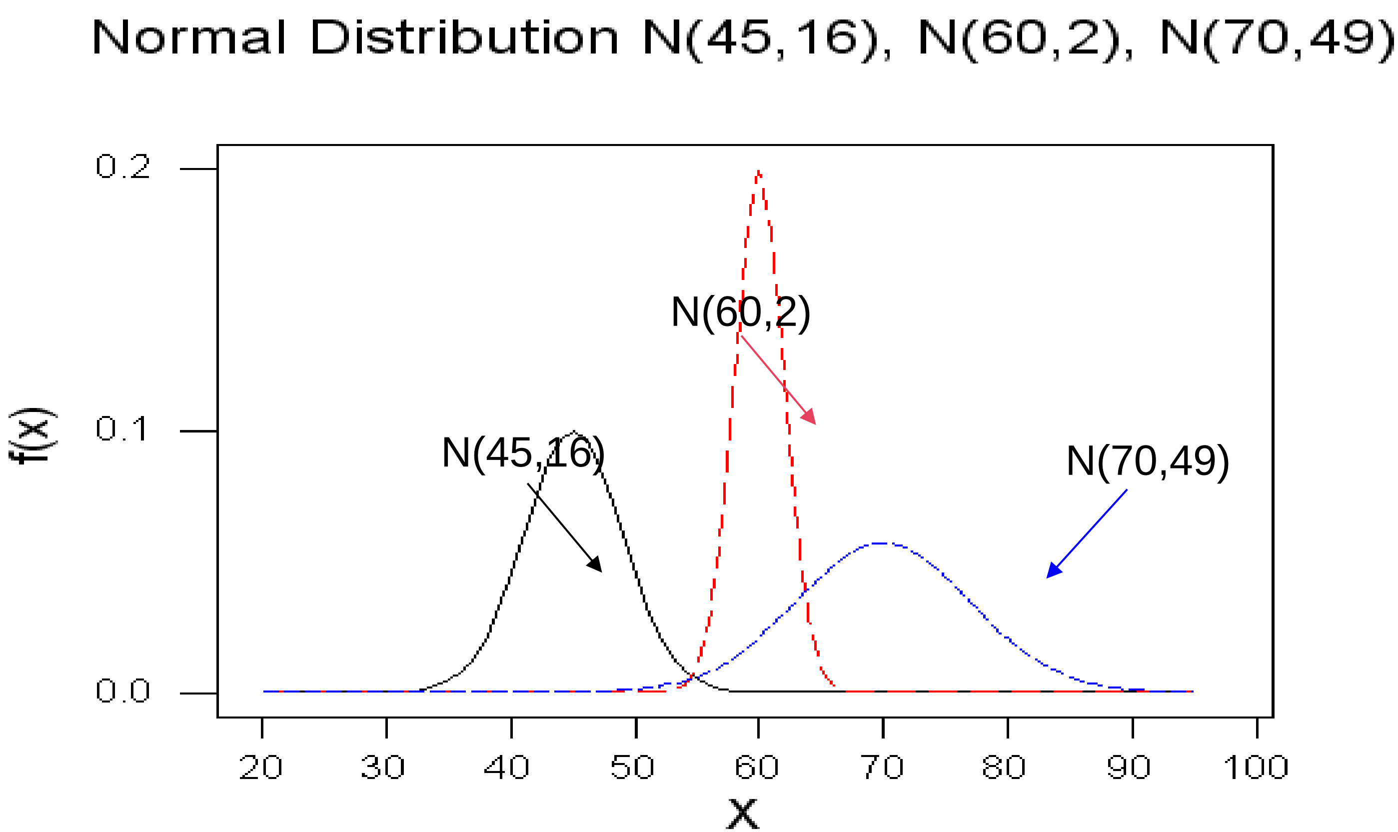
- 평균이 같고 분산이 다른 분포의 형태

Normal Distribution N(55,16), N(55,4), N(55,1)



정규분포의 특징(2)

- 평균과 분산이 다른 분포의 형태



정규분포의 확률밀도함수의 특성

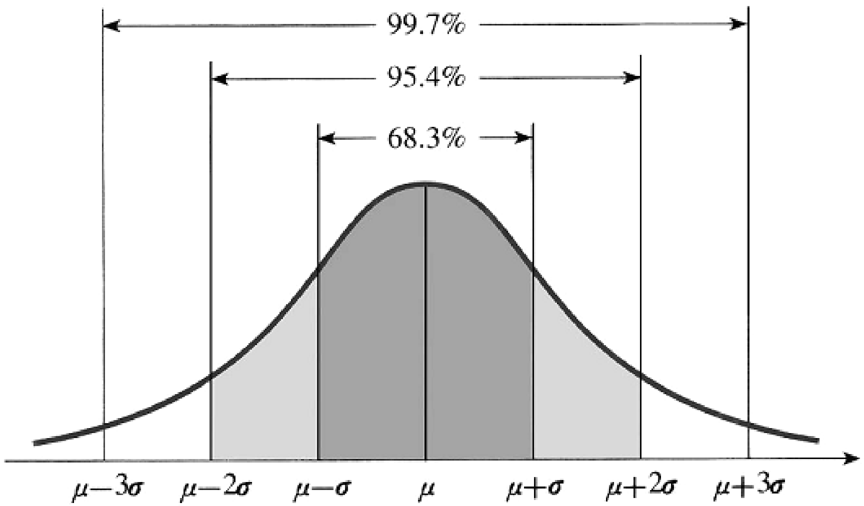


그림 4-8. 정규분포의 확률밀도 함수

1) 정규분포는 평균 중심으로 좌우대칭이며, 종(bell)모양

이 때, 평균(μ), 중앙값, 최빈값 모두 일치

2) 평균으로부터 양쪽으로

구간의 면적이 전체 밀도

비율의 0.683

2배 거리는 전체면적비율의

3배 거리는 0.0970이 되는 특징 있음

$$P(\mu-\sigma \le X \le \mu+\sigma) = 0.683$$

$$P(\mu-2\sigma \le X \le \mu+2\sigma) = 0.954$$

$$P(\mu-3\sigma \le X \le \mu+3\sigma) = 0.997$$

정규분포의 확률밀도함수의 특성

3) 정규곡선 아래의 전체면적은 "1"이다

4) 정규곡선의 모양과 위치는 분포의 평균 과 분산 또는 표준편차에 의해 결정됨

평균은 곡선의 위치를, 분산은 분포의 모양을 나타냄

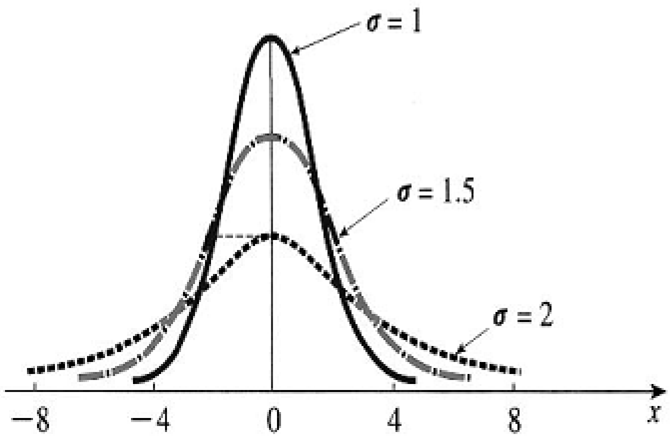


그림 4-9. 평균은 같고, 표준편차가 다른 정규분포 곡선

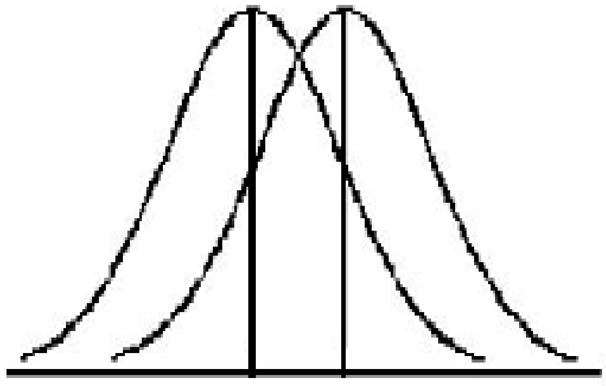


그림 4-10. 평균은 다르지만 표준편차가 같은 정규분포 곡선

표준정규분포

- **표준화(standardization)**란?
- **각기 다른 모양과 위치를 가지는 정규분포들을 하나의 공통된 정규분포로 전환시키는 방법**
- **X값이 평균으로부터 떨어져 있는 거리를 표준편차의 비로 변환하는 것**
- **확률변수 X의 평균이 μ, 표준편차가 σ 일 때**
- **X를** $Z = \dfrac{X_i - \mu}{\sigma}$ **와 같이 변환시키는 것을**

표준화한다고 함

표준정규분포

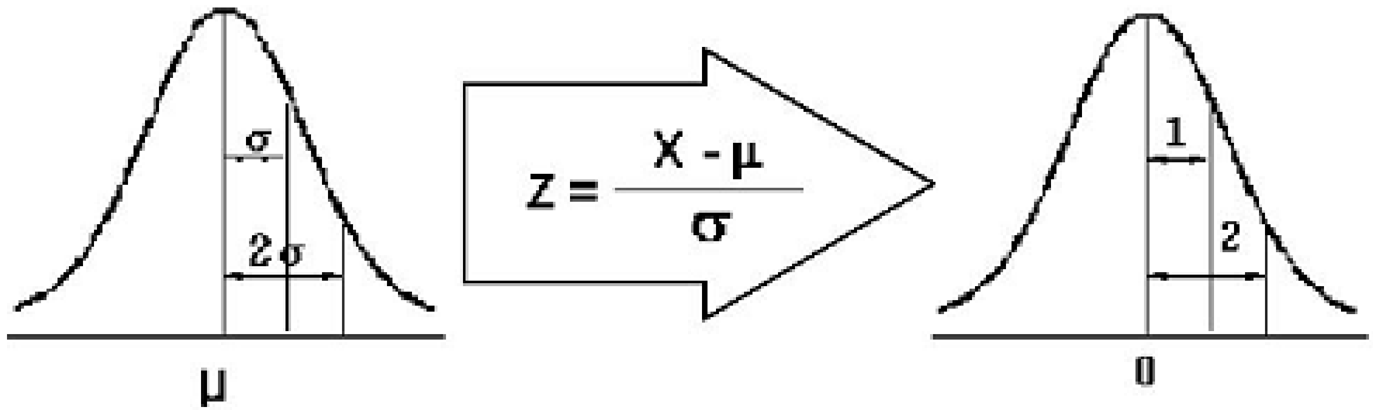


그림 4-11. 정규분포와 표준정규분포와의 관계

- X를 표준화시킨 확률변수 Z의 특성
- Z의 평균 E(Z)는

$$E(Z) = E(\frac{X-\mu}{\sigma}) = E[\frac{X}{\sigma} - \frac{\mu}{\sigma}] = \frac{E(X)}{\sigma} - \frac{\mu}{\sigma} = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$$

- 분산 Var(Z)는

$$Var(Z) = Var\frac{X-\mu}{\sigma}) = Var(\frac{1}{\sigma}(X-\mu)) = \frac{1}{\sigma^2} Var(X-\mu) = \frac{1}{\sigma^2}\sigma^2 = 1$$

표준정규분포

- 이를 정리하면,

$$Z = \frac{X-\mu}{\sigma} \sim N(0,\ 1)$$

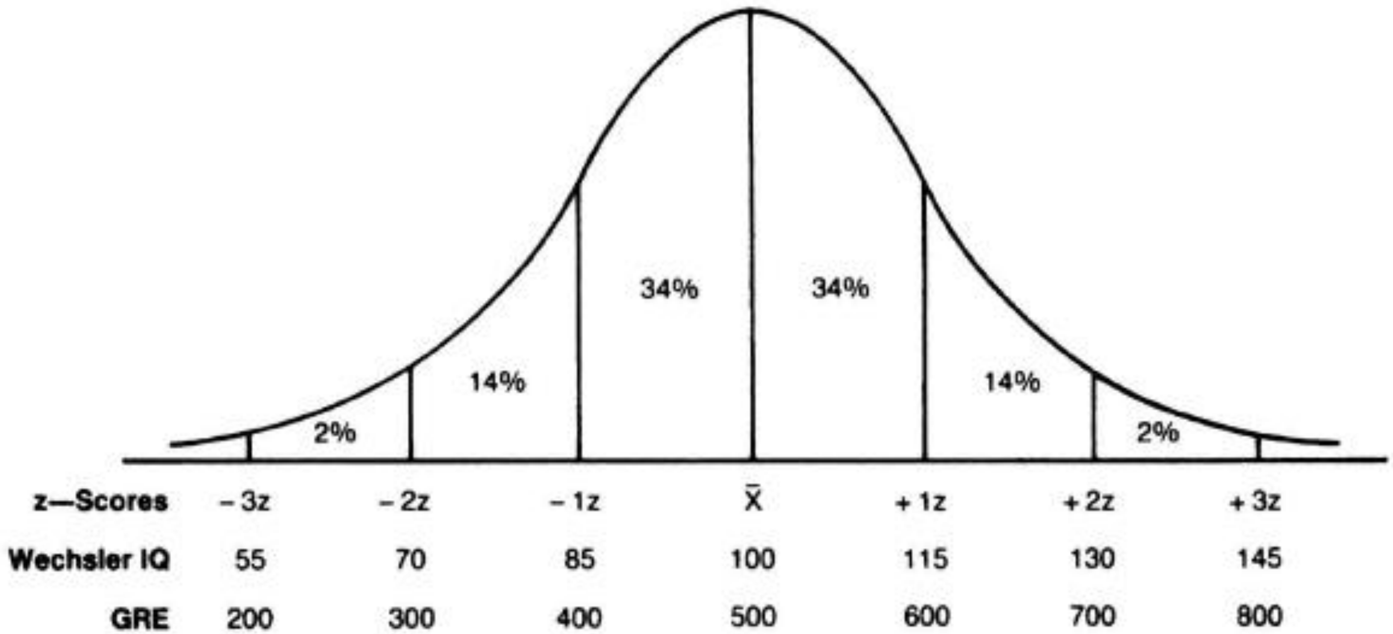
- 특성은 앞에서 설명한 정규분포의 특성과 동일

확률변수 X를 표준화한 Z의 확률밀도함수를 구하기 위하여 정규분포의 확률밀도함수에 (X-μ)/σ 대신 Z를 대입하여 정리하면 다음과 같다

$$f(z) = \frac{1}{\sqrt{2\pi}}\ e^{-\frac{z^2}{2}}, \qquad -\infty < Z < \infty$$

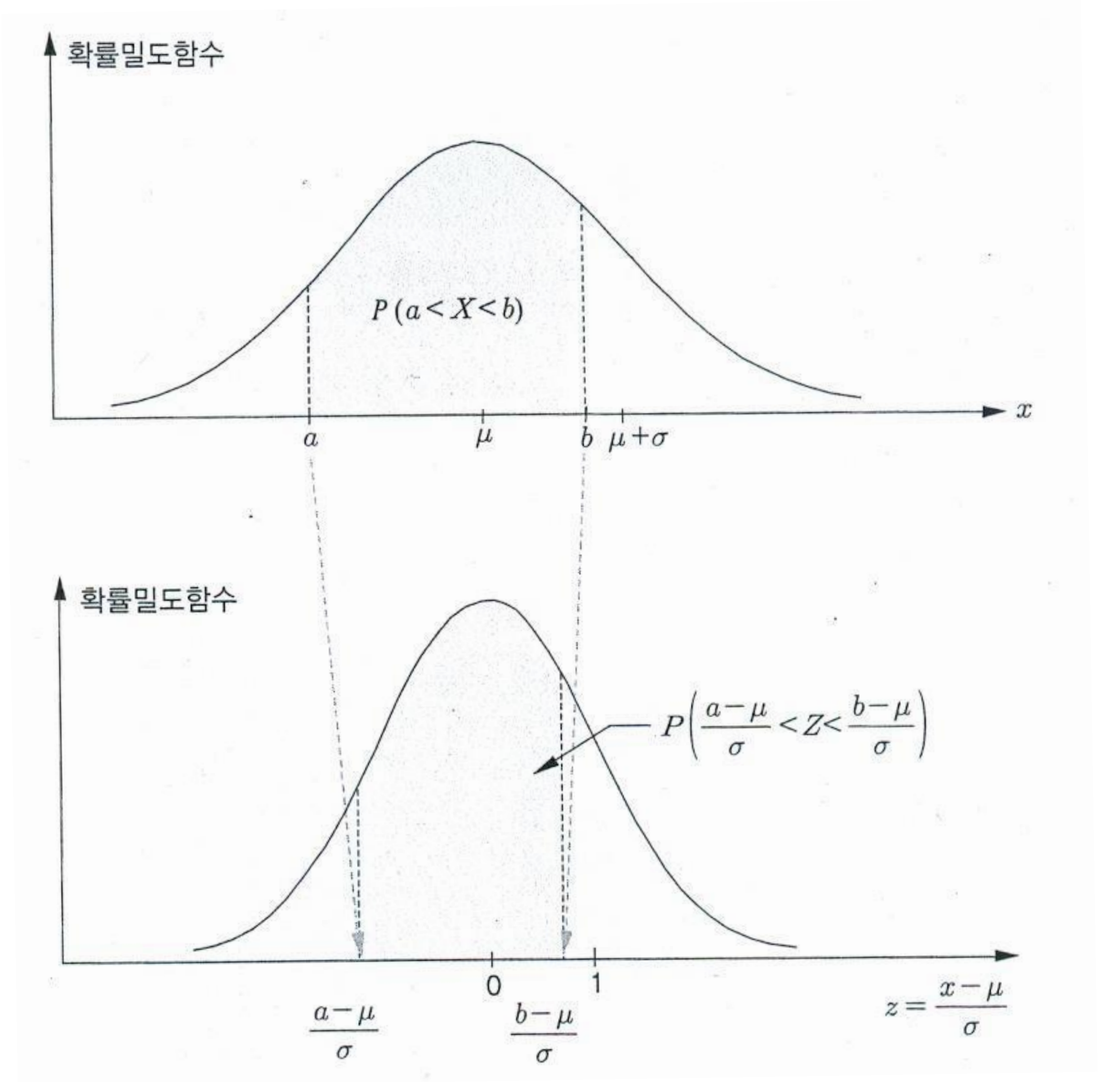
표준정규분포

- 표준화 : z = (점수 - 평균) / 표준편차



*GRE : Graduation Requirement Examination

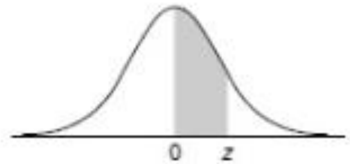
표준 정규분포



확률밀도함수
$P(a<X<b)$
x
a
μ
b
$\mu+\sigma$
확률밀도함수
$P\left(\frac{a-\mu}{\sigma}<Z<\frac{b-\mu}{\sigma}\right)$
0
1
$\frac{a-\mu}{\sigma}$
$\frac{b-\mu}{\sigma}$
$z=\frac{x-\mu}{\sigma}$

Table AIV.2 Standard Norms Table

Area between 0 and z



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998