Lecture Notes for
Introduction to Probability and Statistics

Hanyang University, Ansan, 425-791, Korea
calltimy@hanyang.ac.kr

Probability and Statistics / 확률과 통계

강의노트 14

정규분포

98. 정규분포, Normal Distribution : $N(\mu,\sigma^2)$

Definition 4.4.1 (Normal distribution). A random variable X with density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(1/2)[(x-\mu)/\sigma]^2} \qquad -\infty < x < \infty$$

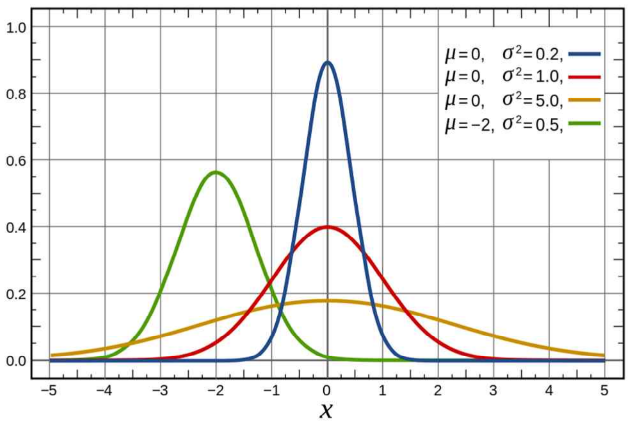
$$-\infty < \mu < \infty$$

$$\sigma > 0$$

is said to have a normal distribution with parameters μ and σ.

99. 표준정규분포, Standard Normal Distribution : N(0,1)

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$



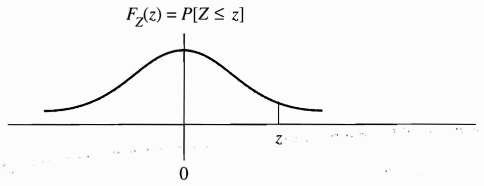
100. z변환 : 평균 μ, 표준편차 σ 인 정규확률분포를 평균 0, 표준편차 1인 표준정규확률변수로 바꿔준다.

$$z = \frac{x-\mu}{\sigma}$$

[참고] 중심극한정리 : Central Limit Theorem (CTL)
평균 μ, 표준편차 σ 인 모집단에서 크기 n 인 표본들을 무작위로 추출하면 n 이 커질수록 $\overline{X}$ 는 평균 μ, 표준편차 $\sigma/\sqrt{n}$ 인 정규분포에 가까워진다.

101. 표준정규분포표 이용

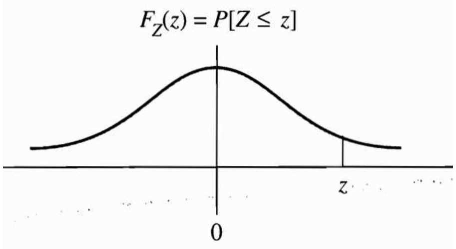
TABLE V
Cumulative distribution: Standard normal



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
	$F_Z(z) = P[Z \le z]$									
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1921	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

>> F(a) = P(z<a) 이고, 이는 그래프에서 z =a 왼쪽의 면적

102. P[a<z<b]



>> P[a<z<b] = F[b] - F[a]

Q. 학생들의 몸무게 평균이 150, 표준편차가 20 이라고 가정하자. 그때 임의의 한 학생을 선택했을 때 그의 몸무게가 170 파운드 이상인 확률은 얼마인가?

A.

103. Theorem 4.5.1

>> P[-σ<X< σ] ≒ .68

>> P[-2σ<X<2σ] ≒ .95

>> P[-3σ<X<3σ] ≒ .997

104. Normal Distribution & Binomial Distribution

$X \sim B(n,p) \rightarrow Z(\mu,\sigma^2)$

$\mu = np$

$\sigma^2 = np(1-p)$

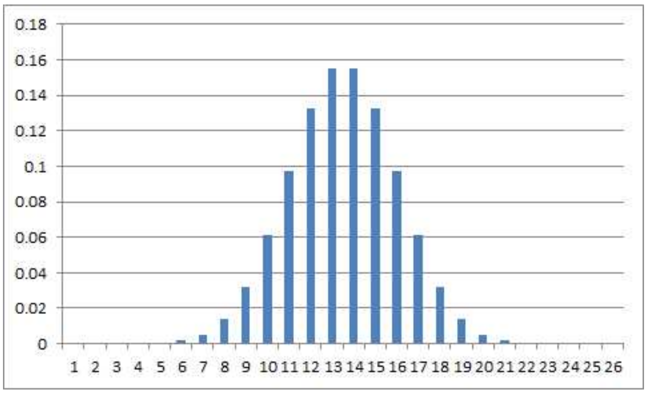
조건 :

$p \leq .5$ 이고 $np > 5$ 이거나

$p > .5$ 이고, $n(1-p) > 5$ 이어야 한다.

105. 25회 동전던지기 이항분포와 정규분포

$X \sim B(25, 0.5)$



1) P[X≤14]을 이항분포로 구해보면,

F(14) = f(0) + f(1) + f(2) + ... + f(14)

= .7878

2) P[X≤14]을 정규분포로 구해보면,

$$\mu = np = 25 \times 0.5 = 12.5$$

$$\sigma^2 = np(1-p) = 25 \times 0.5 \times 0.5 = 6.25$$

$$\sigma = \sqrt{6.25} = 2.5$$

$P[X \leq 14] = P[Z \leq \frac{14-12.5}{2.5}] = P[Z \leq 0.6] = .7256$

조금 더 자세히 살펴보면 P[X≤14] 는 실제로는 X = 14.5 이하의 막대그래프의 면적을 의미한다. 이 .5를 감안하면,

P[X≤14.5] = P[Z≤.8] = .7881

∴ 이항분포로 구한 .7878 과 정규분포로 구한 .7881 은 매우 유사한 근사값으로 이항분포대신 정규분포를 이용하여 계산할 수 있다.