

# 일표본 정규모집단

표준화(Standardization)

$$X \sim N(\mu, \sigma^2) \longrightarrow Z = \frac{X - \mu}{\sigma} \sim N(0, 1) \longrightarrow Z^2 = \left(\frac{X - \mu}{\sigma}\right)^2 \sim \chi^2(1)$$

## 일표본 정규모집단

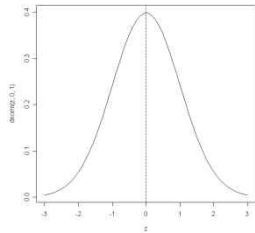
$$X_1, \dots, X_n: \text{RS from } N(\mu, \sigma^2) \longrightarrow Z_1, \dots, Z_n: \text{RS from } N(0, 1) \longrightarrow Z_1^2, \dots, Z_n^2: \text{RS from } \chi^2(1)$$

### 표본평균(Sample Mean)

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

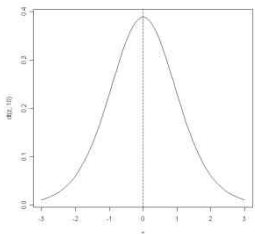
### Standardization

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



$$Z \sim N(0, 1)$$

$$V \sim \chi^2(k) \Rightarrow Z \perp V$$



### 표본분산(Sample Variance)

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$(n-1)s^2 = \sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2$$

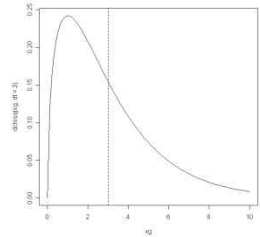
$$\frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 - \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(n-1)$$

### 스튜던트화

### (Studentization)

$$T = \frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)s^2/\sigma^2}{(n-1)}}} = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

$$\sum_{i=1}^n Z_i^2 \sim \chi^2(n)$$



$$V \sim \chi^2(k)$$

$$E[V] = k$$

$$Var[V] = 2k$$

$$x_1 = 1, x_2 = 2, x_3 = 3$$

$$\bar{x} = \frac{1}{3}(1 + 2 + 3) = 2$$

$$s^2 = \frac{1}{3-1} \{(1-2)^2 + (2-2)^2 + (3-2)^2\} = 1$$

$$= \frac{1}{3-1} \{1^2 + 2^2 + 3^2 - 3 \times 2^2\} = 1$$

## 독립표본 정규모집단

독립표본 정규모집단 (이분산)

$X_1, \dots, X_{n1}: RS \text{ from } N(\mu_1, \sigma_1^2)$

$Y_1, \dots, Y_{n2}: RS \text{ from } N(\mu_2, \sigma_2^2)$

$\forall (i, j) X_i, Y_j \text{ independent}$

표본평균

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$$

$$\bar{Y} = \frac{1}{n_2} \sum_{j=1}^{n_2} Y_j \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

표본분산

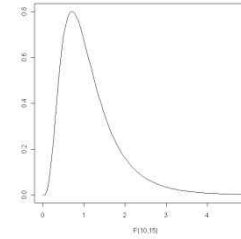
$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$$

$$s_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2$$

$$\frac{(n_1 - 1)s_1^2}{\sigma_1^2} \sim \chi^2(n_1 - 1)$$

$$\frac{(n_2 - 1)s_2^2}{\sigma_2^2} \sim \chi^2(n_2 - 1)$$

$$\frac{\left(\frac{s_1^2}{\sigma_1^2}\right)}{\left(\frac{s_2^2}{\sigma_2^2}\right)} \sim F(n_1 - 1, n_2 - 1)$$



$$V_1 \sim \chi^2(k_1)$$

$$V_2 \sim \chi^2(k_2) \Rightarrow F = \frac{V_1/k_1}{V_2/k_2} \sim F(k_1, k_2)$$

$$V_1 \perp V_2$$

독립표본 정규모집단 (등분산)

$$\sigma_1^2 = \sigma_2^2 = \sigma^2$$

$X_1, \dots, X_{n1}: RS \text{ from } N(\mu_1, \sigma^2)$

$Y_1, \dots, Y_{n2}: RS \text{ from } N(\mu_2, \sigma^2)$

$\forall (i, j) X_i, Y_j \text{ independent}$

표본평균

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right)$$

$$\bar{Y} = \frac{1}{n_2} \sum_{j=1}^{n_2} Y_j \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right)$$

표본분산

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$$

$$s_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2$$

$$\frac{(n_1 - 1)s_1^2}{\sigma^2} \sim \chi^2(n_1 - 1)$$

$$\frac{(n_2 - 1)s_2^2}{\sigma^2} \sim \chi^2(n_2 - 1)$$

합동표본분산(Pooled sample variance)

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{\sigma^2} = \frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2)$$

두 표본평균의 차이의 분포

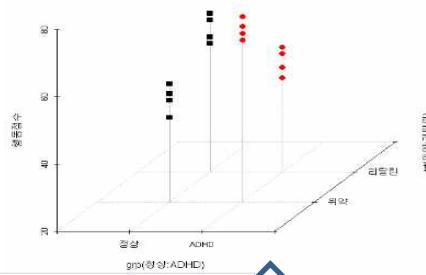
$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

두 표본평균의 차이의  
표준화

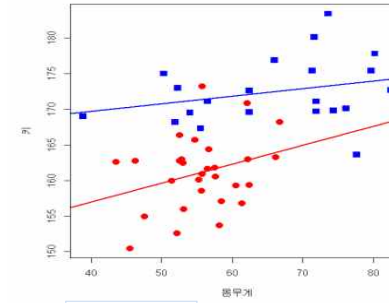
$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0, 1)$$

Studentization

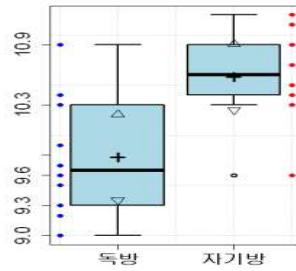
$$T = \frac{\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}{\sqrt{\left(\frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2}\right) / (n_1 + n_2 - 2)}} = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$



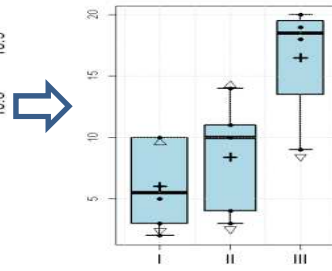
이원 분산분석  
삼원분할표/로그선형모형/로짓모형



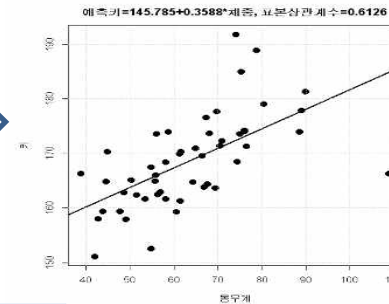
공분산분석  
로짓모형



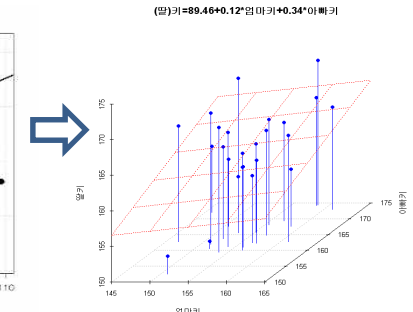
독립표본 t검정/Welch t-검정  
Mann-Whitney U-검정  
Wilcoxon 순위합 검정(Rank Sum)  
이표본 이항정확검정/비율근사검정



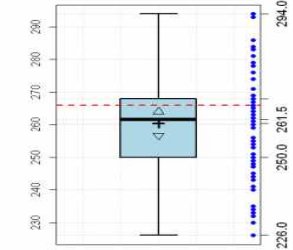
일원 분산분석/Welch, Brown-Forsythe ANOVA  
Kruskal-Wallis 검정  
카이제곱 동질성/독립성 검정



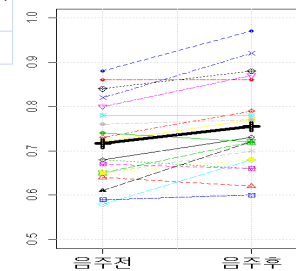
Pearson 상관계수/선형 단순회귀분석  
Spearman 순위상관계수  
로짓모형



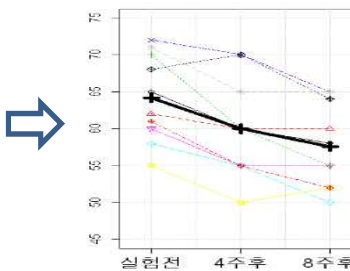
선형 중회귀분석  
로짓모형



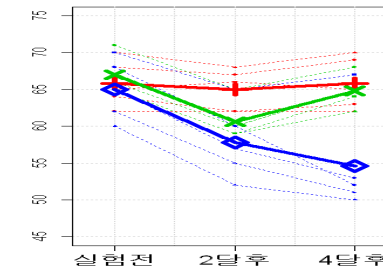
일표본 t검정  
Wilcoxon 부호순위 검정(Signed Rank)  
부호검정 (Sign)  
일표본 이항정확검정/비율 근사검정



대응표본 t검정  
Wilcoxon 부호순위 검정(Signed Rank)  
부호검정(Sign)  
McNemar 주변동질성 검정



반복측정(개체간요인 없음)  
Friedman 검정 (비모수 RCBD)  
Cochran Q검정



반복측정(개체간요인 있음)

## 표준정규분포

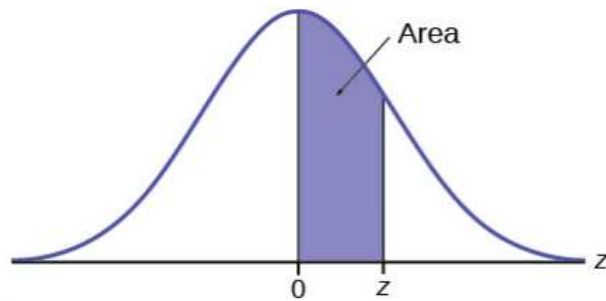
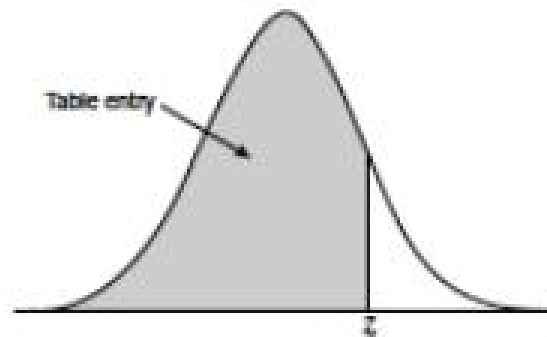
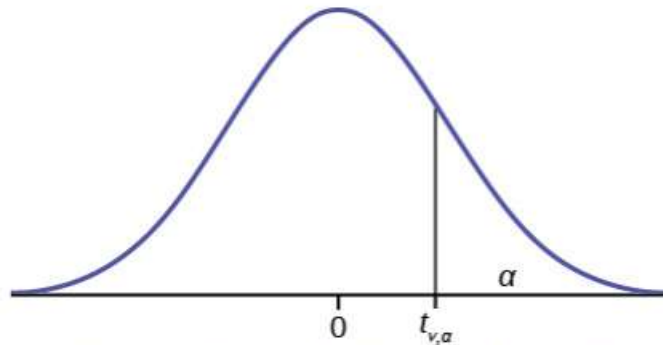


Figure A2



표준정규분포가 1.96보다 작을 확률	표준정규분포의 97.5%백분위수
$P(Z < 1.96) = 0.975$	$z_{0.025} \text{ s.t. } P(Z > z_{0.025}) = 0.025$
<code>pnorm(1.96)</code> <code>[1] 0.9750021</code>	<code>qnorm(0.025, lower=FALSE)</code>
<code>from scipy.stats import norm</code>	
<code>norm.cdf(1.96,0,1)</code> <code>0.9750021048517795</code>	<code>norm.ppf(0.975, 0, 1)</code> <code>1.959963984540054</code>
$P(Z < 1.645) = 0.95$	$z_{0.05} \text{ s.t. } P(Z > z_{0.05}) = 0.05$
<code>pnorm(1.645)</code>	<code>qnorm(0.95)</code>
<code>norm.cdf(1.645,0,1)</code>	<code>norm.ppf(0.95, 0, 1)</code>

[illegible]

Student's  $t$  DistributionFigure A3 Upper critical values of Student's  $t$  Distribution with  $v$  Degrees of Freedom

For selected probabilities,  $\alpha$ , the table shows the values  $t_{v,\alpha}$  such that  $P(t_v > t_{v,\alpha}) = \alpha$ , where  $t_v$  is a Student's  $t$  random variable with  $v$  degrees of freedom. For example, the probability is .10 that a Student's  $t$  random variable with 10 degrees of freedom exceeds 1.372.

자유도가 4인 t분포의 99% 백분위수

$$t_{0.01}^{(4)} \text{ s.t. } P(T(4) > t_{0.01}^{(4)}) = 0.01$$

```
qt(0.01, df=4, lower=FALSE)
[1] 3.746947
```

```
from scipy.stats import t
t.ppf(0.99, 4)
3.7469473879811366
```

df	0.250	0.100	0.050	0.025	0.010	0.005
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.695	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.624	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.319	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
40	0.681	1.303	1.684	2.021	2.423	2.704
60	0.679	1.296	1.671	2.000	2.390	2.660
120	0.677	1.289	1.658	1.980	2.358	2.617
Inf	0.674	1.282	1.645	1.960	2.326	2.576

# 카이제곱분포 표

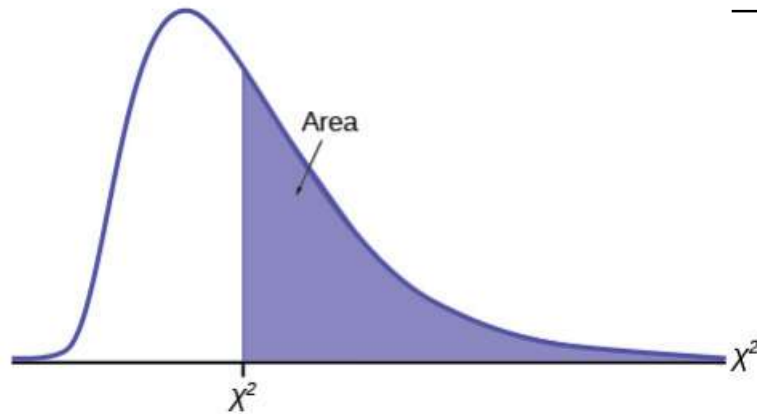


Figure A4

자유도가 4인 카이제곱분포의 95%백분위수

$$\chi^2_{0.01}(4) \text{ s.t. } P(X^2 > \chi^2_{0.01}(4)) = 0.01$$

```
qchisq(0.05, df=4, lower=FALSE)
[1] 9.487729
```

```
from scipy.stats import chi2
chi2.ppf(0.95, 4)
9.487729036781154
```

df	0.995	0.990	0.975	0.950	0.050	0.025	0.010	0.005
1	0.0000	0.0002	0.0010	0.0039	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	5.9915	7.3778	9.2103	10.5966
3	0.0717	0.1148	0.2158	0.3519	7.8147	9.3484	11.3449	12.8382
4	0.2070	0.2971	0.4844	0.7107	9.4877	11.1433	13.2767	14.8603
5	0.4117	0.5543	0.8312	1.1455	11.0705	12.8325	15.0863	16.7496
6	0.6757	0.8721	1.2373	1.6354	12.5916	14.4494	16.8119	18.5476
7	0.9893	1.2390	1.6899	2.1674	14.0671	16.0128	18.4753	20.2777
8	1.3444	1.6465	2.1797	2.7326	15.5073	17.5346	20.0902	21.9550
9	1.7349	2.0879	2.7004	3.3251	16.9190	19.0228	21.6660	23.5894
10	2.1559	2.5582	3.2470	3.9403	18.3070	20.4832	23.2093	25.1882
11	2.6032	3.0535	3.8158	4.5748	19.6751	21.9201	24.7250	26.7569
12	3.0738	3.5706	4.4038	5.2260	21.0261	23.3367	26.2170	28.2995
13	3.5650	4.1069	5.0088	5.8919	22.3620	24.7356	27.6883	29.8195
14	4.0747	4.6604	5.6287	6.5706	23.6848	26.1190	29.1412	31.3194
15	4.6009	5.2294	6.2621	7.2609	24.9958	27.4884	30.5779	32.8013
16	5.1422	5.8122	6.9077	7.9617	26.2962	28.8454	31.9999	34.2672
17	5.6972	6.4078	7.5642	8.6718	27.5871	30.1910	33.4087	35.7185
18	6.2648	7.0149	8.2308	9.3905	28.8693	31.5264	34.8053	37.1565
19	6.8440	7.6327	8.9065	10.1170	30.1435	32.8523	36.1909	38.5823
20	7.4338	8.2604	9.5908	10.8508	31.4104	34.1696	37.5662	39.9969
21	8.0337	8.8972	10.2829	11.5913	32.6706	35.4789	38.9322	41.4011
22	8.6427	9.5425	10.9823	12.3380	33.9244	36.7807	40.2894	42.7957
23	9.2604	10.1957	11.6886	13.0905	35.1725	38.0756	41.6384	44.1813
24	9.8862	10.8564	12.4012	13.8484	36.4150	39.3641	42.9798	45.5585
25	10.5197	11.5240	13.1197	14.6114	37.6525	40.6465	44.3141	46.9279
26	11.1602	12.1982	13.8439	15.3792	38.8851	41.9232	45.6417	48.2899
27	11.8076	12.8785	14.5734	16.1514	40.1133	43.1945	46.9629	49.6449
28	12.4613	13.5647	15.3079	16.9279	41.3371	44.4608	48.2782	50.9934
29	13.1212	14.2565	16.0471	17.7084	42.5570	45.7223	49.5879	52.3356
30	13.7867	14.9535	16.7908	18.4927	43.7730	46.9792	50.8922	53.6720
40	20.7065	22.1643	24.4330	26.5093	55.7585	59.3417	63.6907	66.7660
50	27.9908	29.7067	32.3574	34.7643	67.5048	71.4202	76.1539	79.4900
60	35.5345	37.4849	40.4818	43.1880	79.0819	83.2977	88.3794	91.9517
70	43.2752	45.4417	48.7576	51.7393	90.5312	95.0232	100.4252	104.2149
80	51.1719	53.5401	57.1532	60.3915	101.8795	106.6286	112.3288	116.3211
90	59.1963	61.7541	65.6466	69.1260	113.1453	118.1359	124.1163	128.2989
100	67.3276	70.0649	74.2219	77.9295	124.3421	129.5612	135.8067	140.1695

# F분포 ( $\alpha=0.05$ )

F Distribution

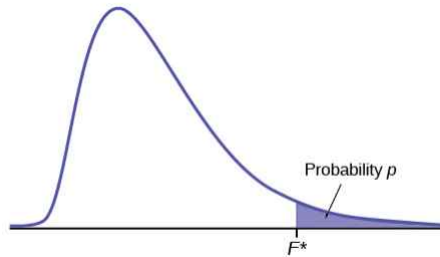


Figure A1 Table entry for  $p$  is the critical value  $F^*$  with probability  $p$  lying to its right.

k1=3, k2=4인 F분포의 95%백분위수

$f(3,4)_{0.05}$  s.t.  $P(F(3,4) > f(3,4)_{0.05}) = 0.05$

`qf(0.05,df1=3,df2=4,lower=FALSE)`  
[1] 6.591382

```
from scipy.stats import f
f.ppf(0.95, 3, 4)
6.591382116425578
```

df2/df1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	Inf
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
Inf	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00