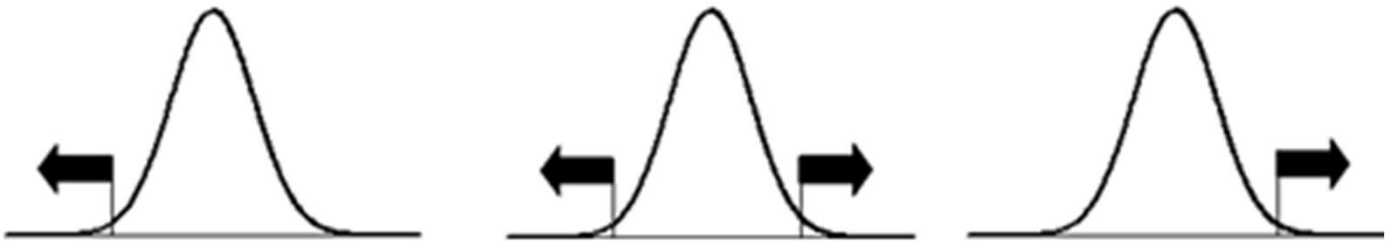
단측검정과 양측검정

양측검정과 단측검정



좌측검정(leftsided test)	양측검정(two-sided test)	우측검정(right-sided test)
$H_0 : p \geq 0.01$	$H_0 : p = 0.01$	$H_0 : p \leq 0.01$
$H_1 : p < 0.01$	$H_1 : p \neq 0.01$	$H_1 : p > 0.01$

대수의 법칙(LAW OF LARGE NUMBER)

◆ 동전을 10번 던져서 앞면과 뒷면이 5번씩 나올 확률?

엑셀 활용

n 번의 베르누이 시행에서 k 번 성공에 대한 경우의 수를 구할 때,

$$=\text{COMBIN}(n,k)=\frac{n!}{k!(n-k)!}, k=0,1,\ldots,n$$

흔히 라고 표기하며, "*n choose k*" 또는 "*n combination k*" 라고 읽는다.

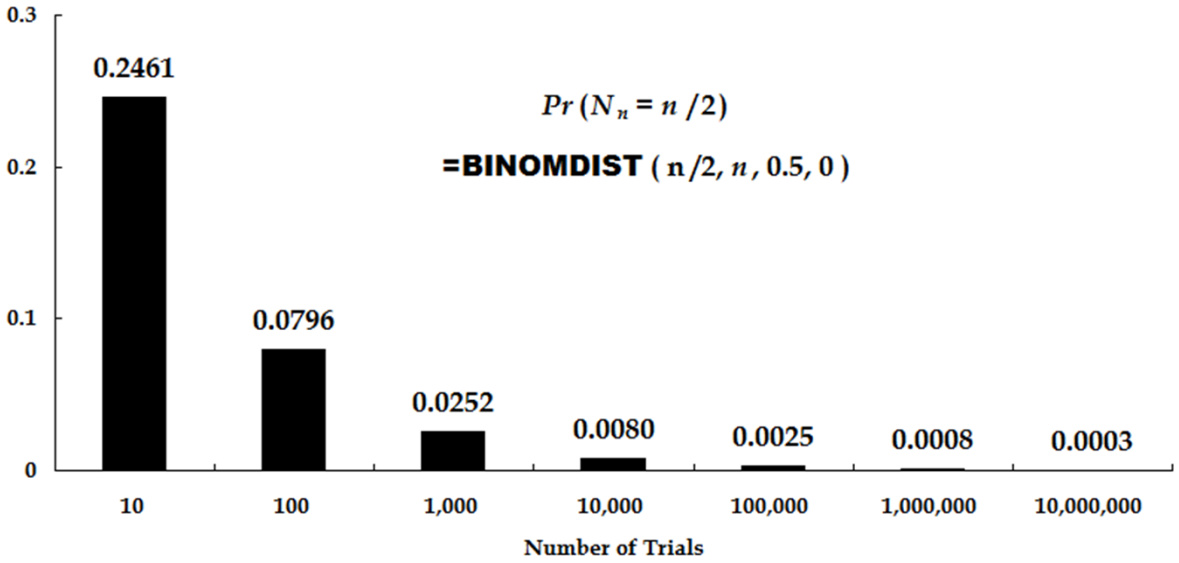
fx =COMBIN(10, H$2)

B	C	D	E	F	G	H	I	J	K	L	M	N
k	0	1	2	3	4	5	6	7	8	9	10	sum
${}_nC_k$	1	10	45	120	210	252	210	120	45	10	1	1024
$Pr(N_n=k)$	0.0010	0.0098	0.0439	0.1172	0.2051	0.2461	0.2051	0.1172	0.0439	0.0098	0.0010	1

$$Pr(N_{10}=5)=\frac{\text{앞면이 5회 나오는 경우의 수}}{\text{총 경우의 수}}=\frac{{}_{10}C_5}{2^{10}}=0.2461=\textbf{BINOMDIST}(5,\ 10,\ 0.5,\ 0)$$

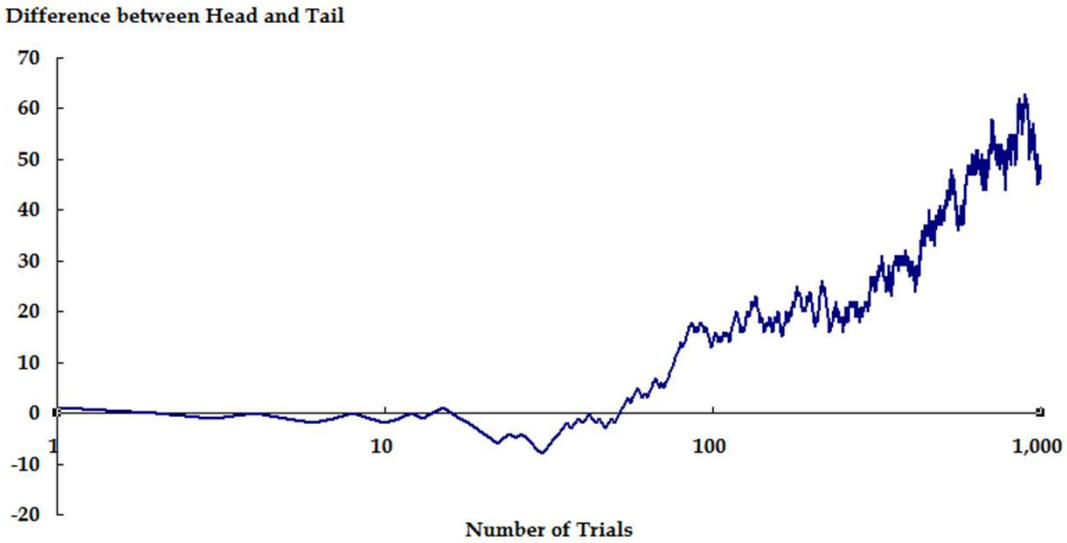
대수의 법칙(LAW OF LARGE NUMBER)

◆ <그림 7-1> 동전던지기 시행횟수와 양면이 동일하게 나올 확률



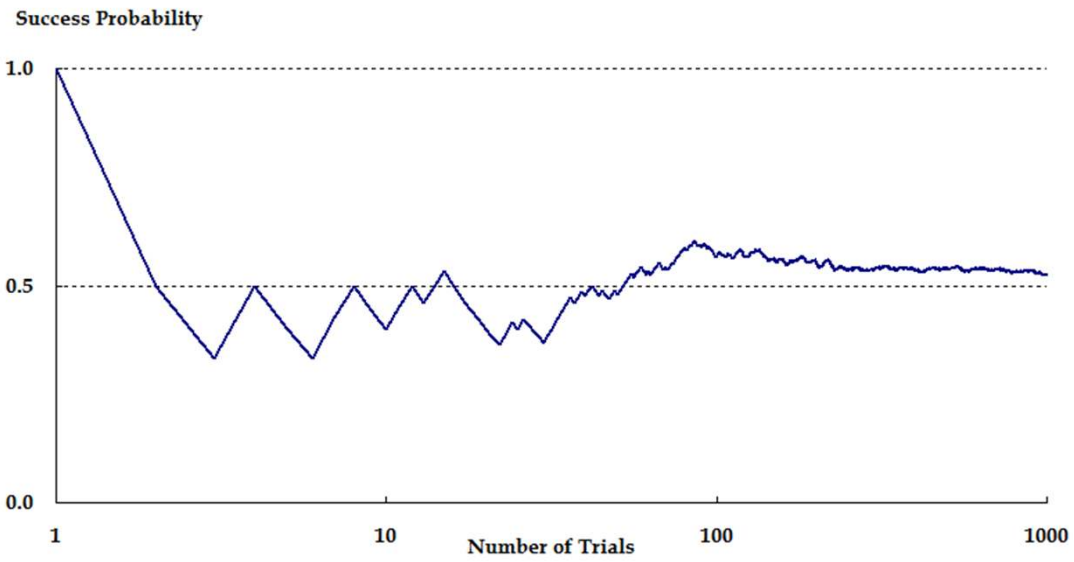
대수의 법칙(LAW OF LARGE NUMBER)

◆ <그림 7-2> 동전던지기 시행횟수와 양면의 관찰빈도의 상대적 차이



대수의 법칙(LAW OF LARGE NUMBER)

◆ <그림 7-3> 동전던지기 시행횟수와 앞면과 뒷면의 관찰빈도의 차이

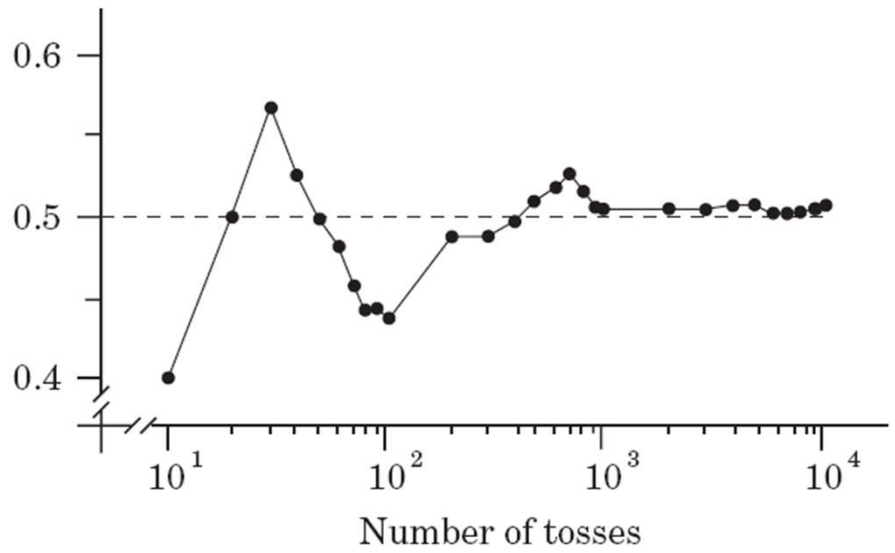


대수의 법칙(LAW OF LARGE NUMBER)

대수의 법칙

동전 던지기를 반복하는 것 같은 확률과정에서, 시행횟수가 늘어남에 따라 관찰확률 혹은 경험적 확률은 이론적 확률에 접근한다는 것이다. 만일 반복의 횟수를 무한대로 증가시키면 이론적 확률에 수렴한다.

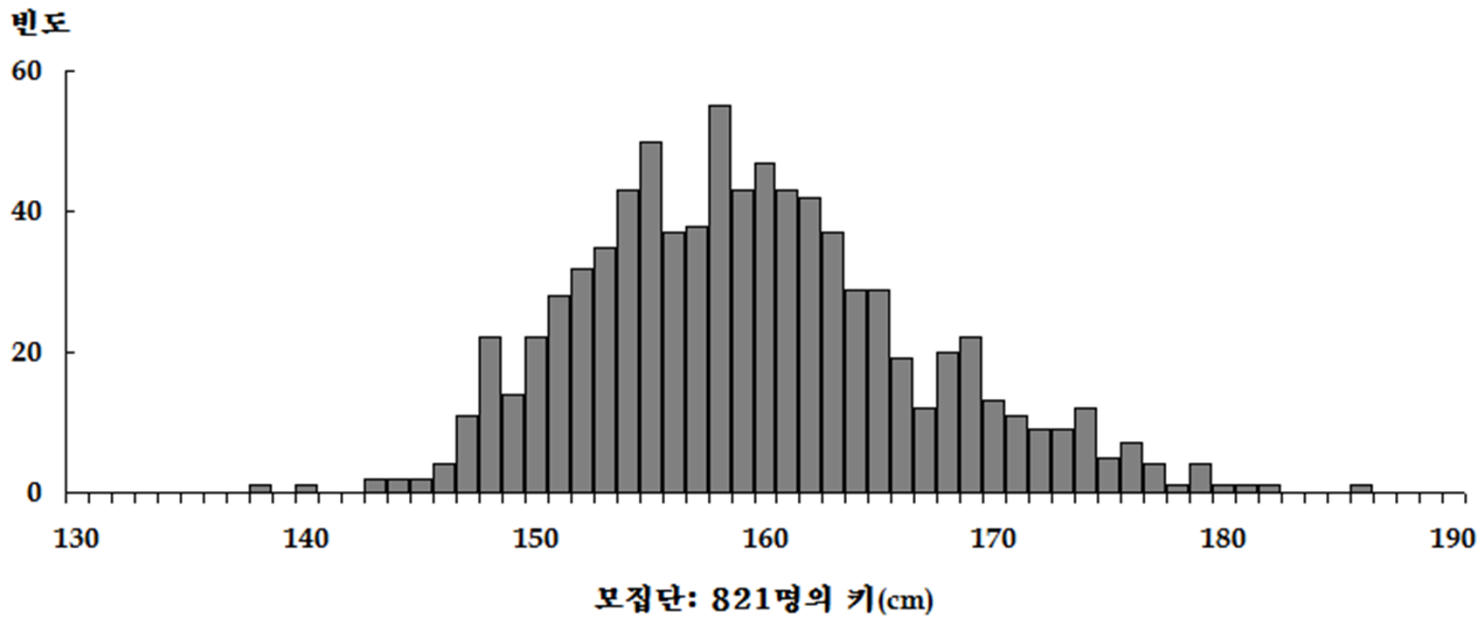
- **Proportion of heads (John Kerrich's coin-tossing experiment)**



표본변동과 표본오차

◆ **모집단 821명의 키(cm)**

●평균(μ)=*159.37* 이고 표준편차(σ)=7.29



표본변동과 표본오차

◆ 모집단에서 10개씩 추출된 10개의 표본

Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10
163	159	164	164	146	163	169	159	159	158
158	161	165	158	167	156	179	149	167	172
158	159	143	155	161	169	165	158	156	152
151	157	162	164	163	148	158	151	153	176
172	153	154	148	163	146	161	163	166	158
167	165	149	147	154	168	160	164	158	149
162	153	163	167	151	164	158	155	161	169
161	176	165	155	154	155	176	140	158	162
163	170	161	148	147	152	164	174	160	154
170	168	165	165	174	157	174	158	159	157
$\overline{X}_1$ =162.5	$\overline{X}_2$ =162.1	$\overline{X}_3$ =159.1	$\overline{X}_4$ =157.1	$\overline{X}_5$ =158	$\overline{X}_6$ =157.8	$\overline{X}_7$ =166.4	$\overline{X}_8$ =157.1	$\overline{X}_9$ =159.7	$\overline{X}_{10}$ =160.7
s=6.169	s=7.534	s=7.767	s=7.695	s=9.080	s=7.997	s=7.706	s=9.267	s=4.218	s=8.932

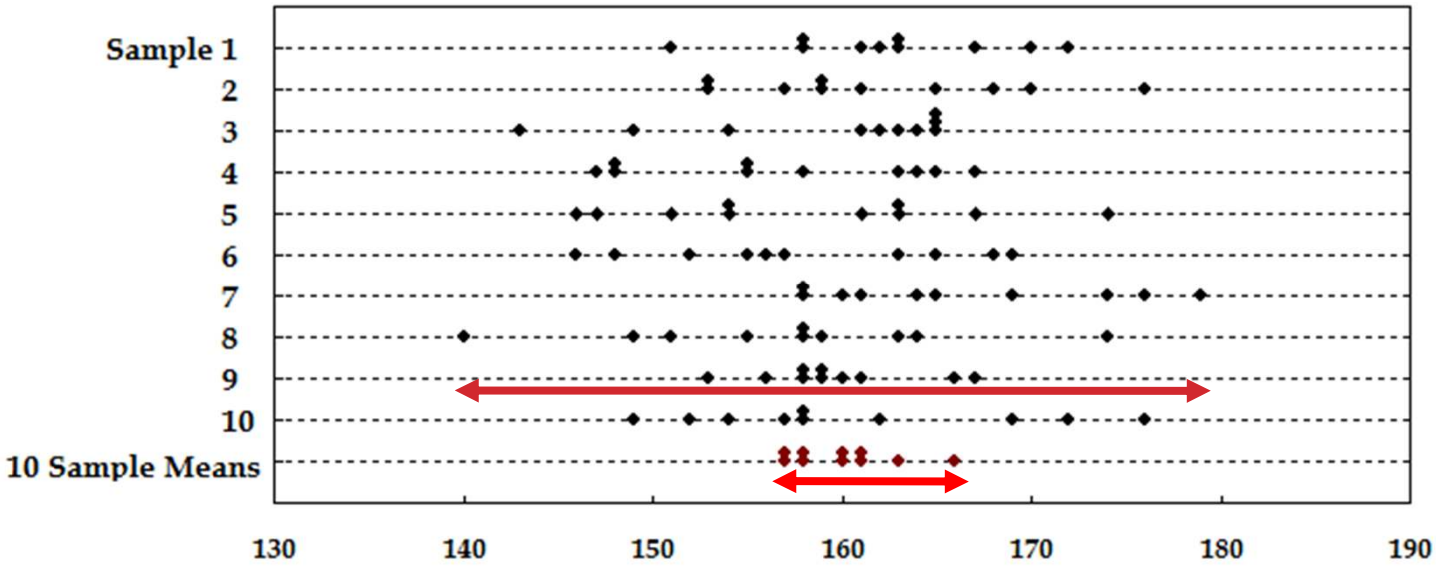
표본변동과 표본오차

◆ 모집단에서 10개씩 추출된 10개의 표본

Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10
163	159	164	164	146	163	169	159	159	158
158	161	165	158	167	156	179	149	167	172
158	159	143	155	161	169	165	158	156	152
151	157	162	164	163	148	158	151	153	176
172	153	154	148	163	146	161	163	166	158
167	165	149	147	154	168	160	164	158	149
162	153	163	167	151	164	158	155	161	169
161	176	165	155	154	155	176	140	158	162
163	170	161	148	147	152	164	174	160	154
170	168	165	165	174	157	174	158	159	157
$\overline{X}_1$ =162.5	$\overline{X}_2$ =162.1	$\overline{X}_3$ =159.1	$\overline{X}_4$ =157.1	$\overline{X}_5$ =158	$\overline{X}_6$ =157.8	$\overline{X}_7$ =166.4	$\overline{X}_8$ =157.1	$\overline{X}_9$ =159.7	$\overline{X}_{10}$ =160.7
s=6.169	s=7.534	s=7.767	s=7.695	s=9.080	s=7.997	s=7.706	s=9.267	s=4.218	s=8.932

표본변동과 표본오차

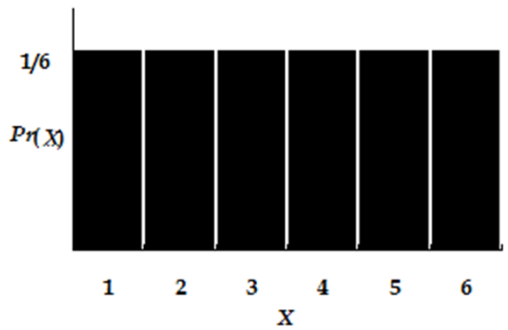
◆ 표본변동과 표본평균의 변동



표본변동과 표본오차

◆ 주사위 던지기(1회 던져 나온 수, n=1)

X	1	2	3	4	5	6
$Pr(X)$	1/6	1/6	1/6	1/6	1/6	1/6

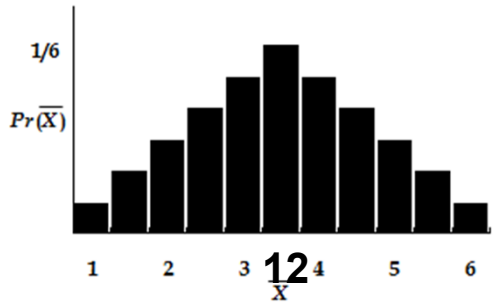


◆ 주사위 던지기(2회 던져 나온 수의 합, n=2)

X_1+X_2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

◆ 표본크기 n = 2인 표본합과 표본평균의 표본분포

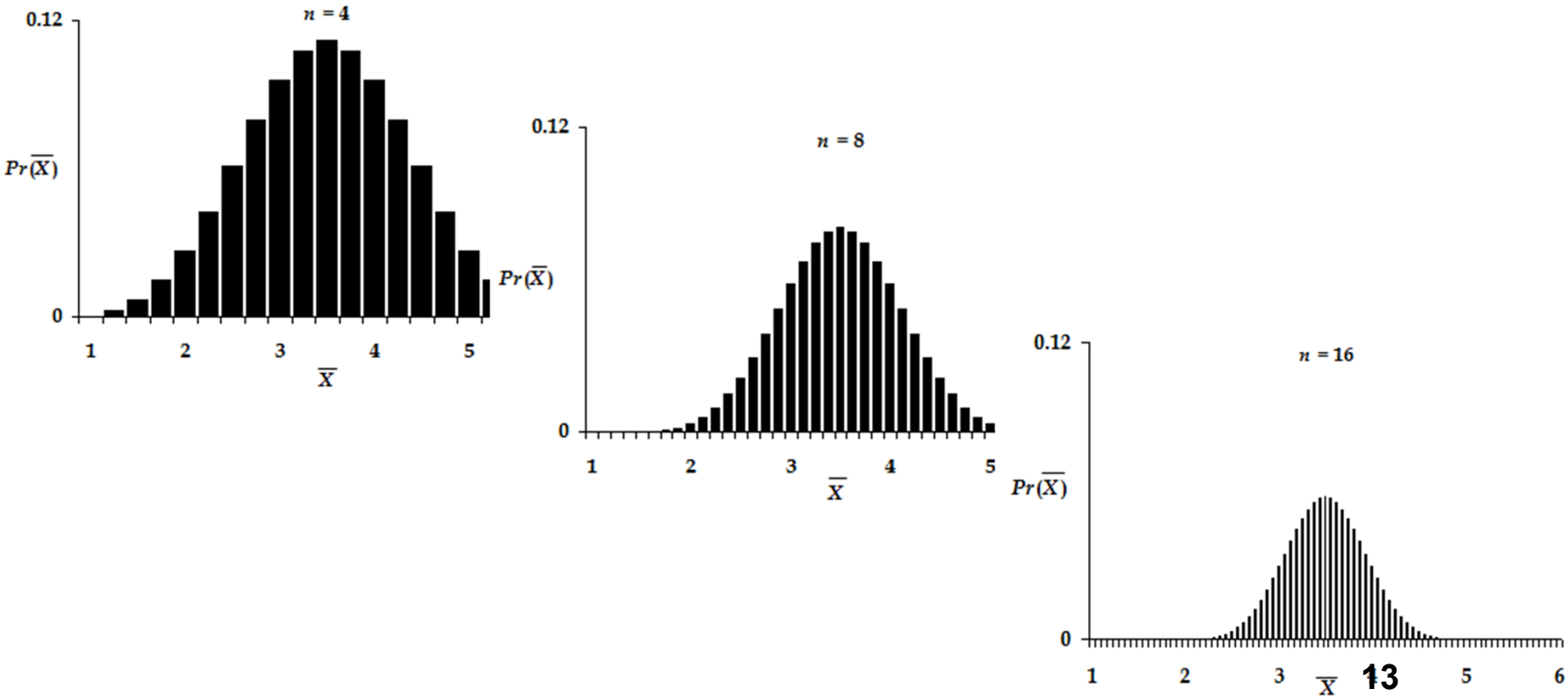
X_1+X_2	2	3	4	5	6	7	8	9	10	11	12
$\overline{X}$	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
$Pr(\overline{X})$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36



중심극한정리(CENTRAL LIMIT THEOREM)

중심극한정리

평균이 μ 이고 분산이 σ^2 인 모집단으로부터 추출한 크기가 n 인 확률표본의 표본평균 $\overline{X}$ 는 n 이 증가할수록 모집단의 분포유형에 상관없이 근사적으로 **정규분포 $N(\mu, \sigma^2/n)$** 을 따른다.



단측검정

동전던지기로 돈 내기를 했다고 하자. 여러 판을 거듭한 끝에 돈을 땄다면, 동전을 문제 삼을 리 없다. 하지만 돈을 왕창 잃었다면 얘기가 다르다. 놀음판에 무엇인가 조작이 있었다고 의심할 것이다.
뒷면이 지나치게 많이 나오는 바람에 돈을 잃었다면, 앞면이 조금 나오도록 조작된 동전이 아닌지를 검정하고 싶을 것이다. 당신의 연구가설은 "$H_1 : p < 0.5$"이다. 부등호가 왼쪽을 째려보고 있으니, **좌측검정**(left-sided test)이다. 반대로 앞면이 지나치게 많이 나와 돈을 잃었다면, 당신의 연구가설은 "$H_1 : p > 0.5$"이다. 이번에는 부등호가 오른쪽을 노려보고 있으니, **우측검정**(right-sided test)이다.

연구 문제 7-1

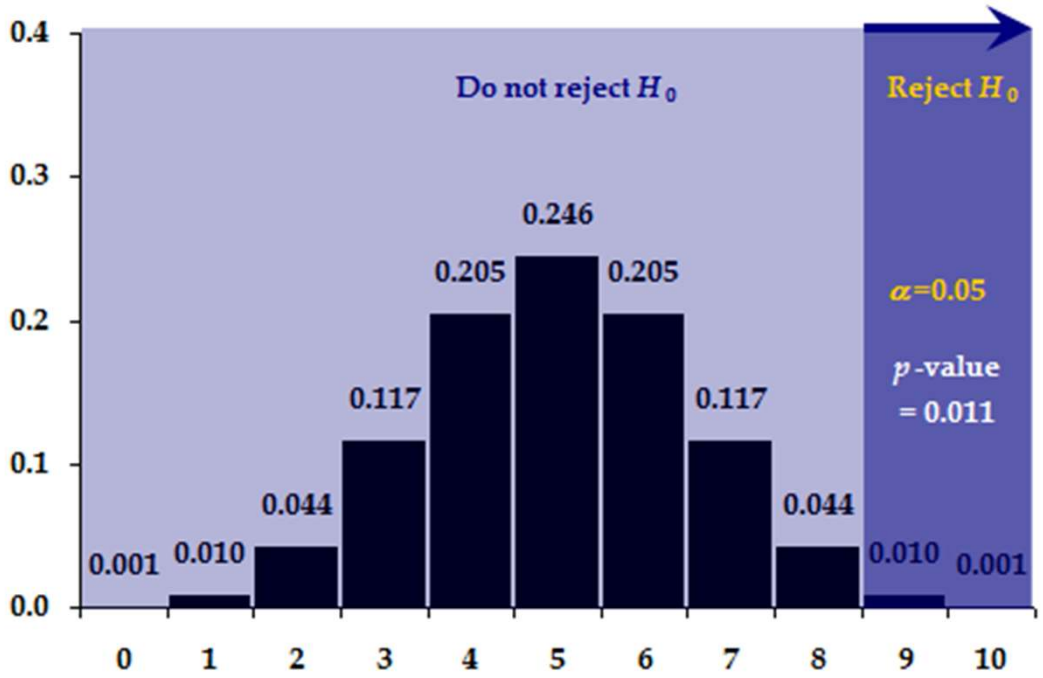
동전을 10번 던졌을 때 앞면이 최소한 몇 번 정도는 나와야, 앞면이 많이 나오는 동전이라고 주장하겠는가? 유의수준 0.05 에서 그 기준을 도출하라.

$H_0 : p \le 0.5$ $\qquad$ $H_1 : p > 0.5$

단측검정

$N_{10}=k$	0	1	2	3	4	5	6	7	8	9	10
$Pr(N_{10}=k)$	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001
p-value(k)*	1	0.999	0.989	0.945	0.828	0.623	0.377	0.172	0.055	0.011	0.001

* p-value(k) = $Pr(N_{10} \geq k \mid \pi = 0.5)$ **= 1−BINOMDIST** (k−1, 10, 0.5, **1**)



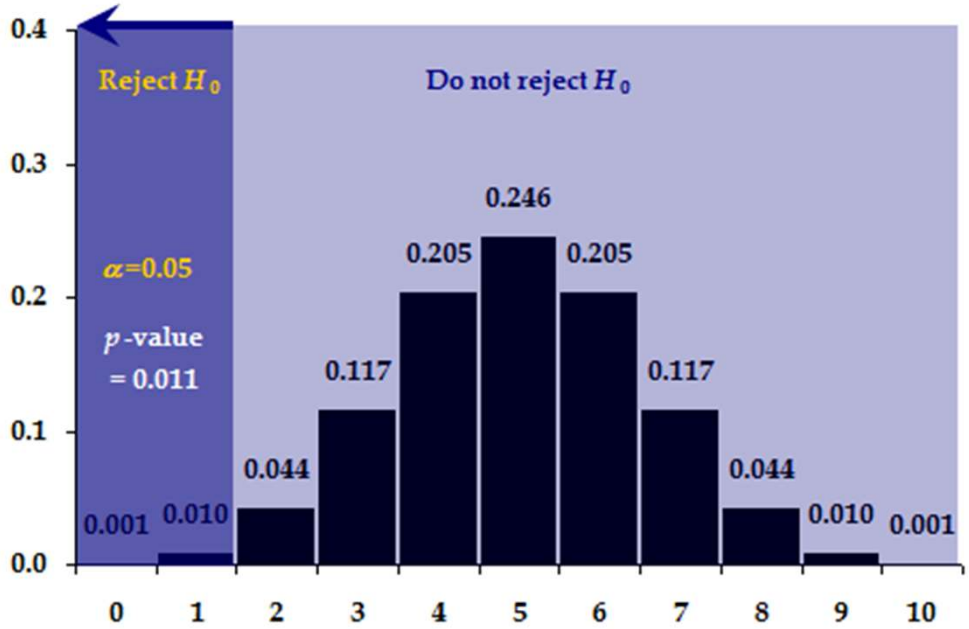
단측검정

연구 문제 7-2

뒷면이 얼마나 많이 관찰될 때, 당신은 앞면보다는 뒷면이 많이 나오는 동전이라고 주장하겠는가? 유의수준 0.05에서 그 기준을 도출하라. $H_0 : p \geq 0.5$ $H_1 : p < 0.5$

$N_{10}=k$	0	1	2	3	4	5	6	7	8	9	10
$Pr(N_{10}=k)$	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001
p-value(k)*	0.001	0.011	0.055	0.172	0.377	0.623	0.828	0.945	0.989	0.999	1

* p-value(k) = $Pr(N_{10} \leq k \mid p = 0.5)$ = BINOMDIST (k, 10, 0.5, 1)

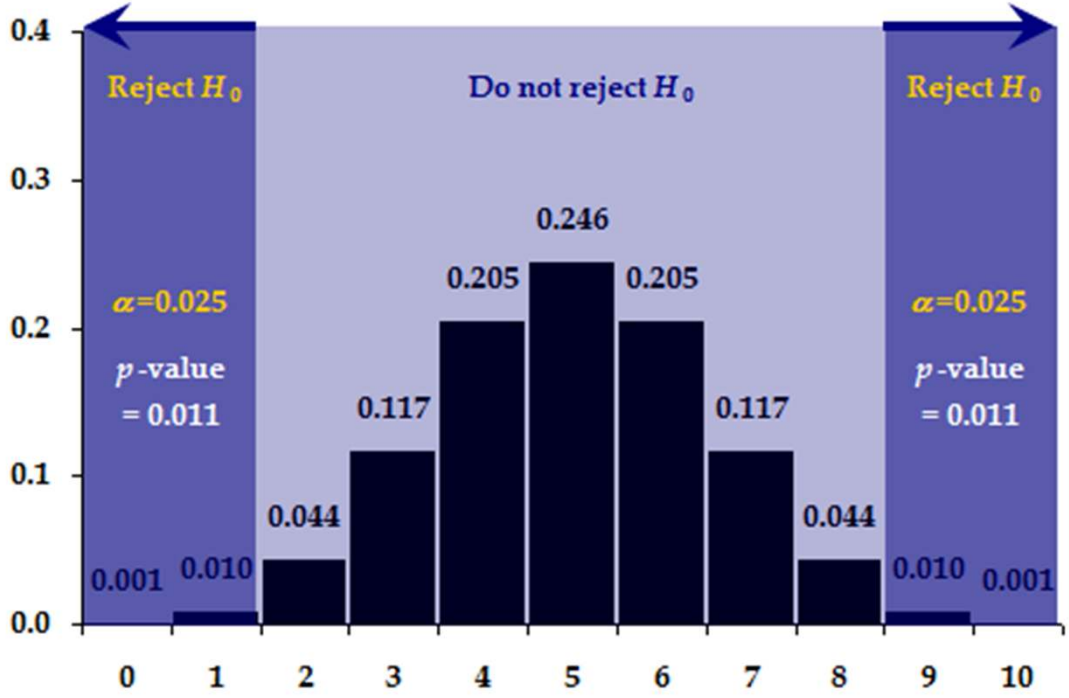


양측검정

연구 문제 7-3

동전던지기 10번에서 앞면과 뒷면의 차이가 얼마나 벌어져야, 당신은 그 동전이 균형 잡힌 동전이 아니라고 주장하겠는가? 즉 특정한 면이 얼마나 많이 관찰되었을 때, 이 동전의 앞면과 뒷면이 나올 확률인 반반이 아니라고 말하겠는가? 유의수준 0.05에서 그 기준을 도출하라.

$H_0 : p = 0.5 \quad H_1 : p \neq 0.5$



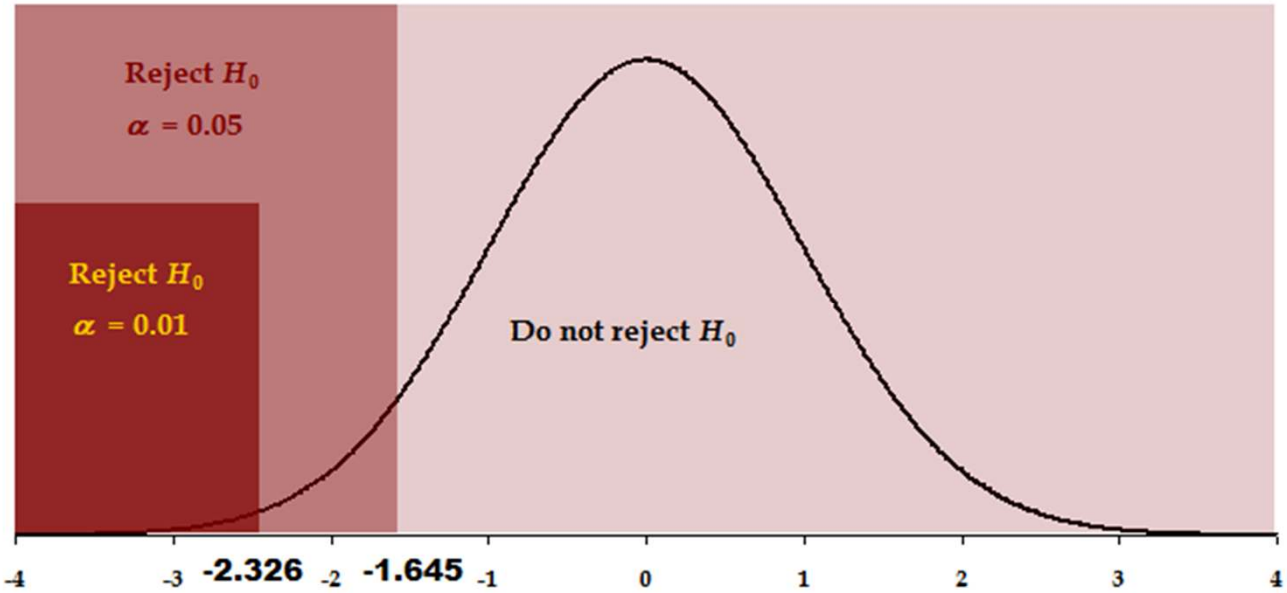
단측검정과 기각역

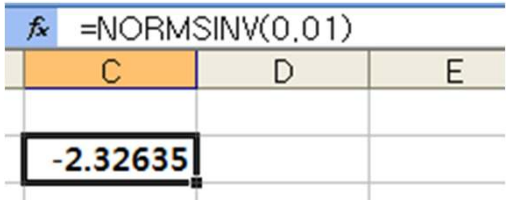
연구 문제 7-4

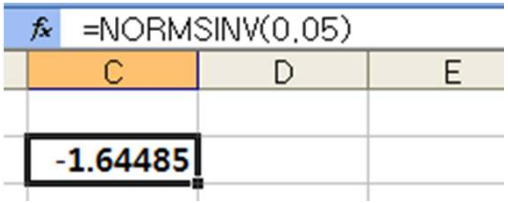
이제 동전 던지기를 100번 수행한다고 하자. 그러면 당신은 관찰결과가 어떤 경우에 앞면이 많이 나오는 동전이라고 주장하겠는가? 적절한 가설을 세우고, 유의수준 0.05에서 그 기준을 도출하라.

$N_{100} = k$	56	57	58	59	60	61	62	63	64	65	66
p-value (k)	0.136	0.097	0.067	0.044	0.028	0.018	0.010	0.006	0.003	0.002	0.001
$N_{100} = k$	67	68	69	70	71	72	73	74	75	76	77
p-value (k)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

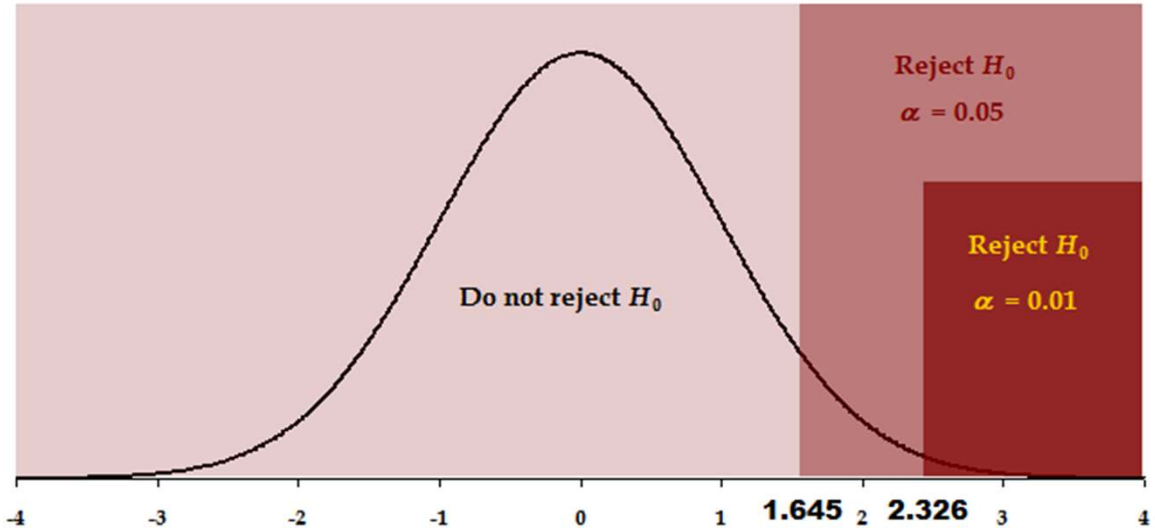
Z-좌측검정과 기각역


Reject H_0
α = 0.05
Reject H_0
α = 0.01
Do not reject H_0
-4
-3
-2.326
-2
-1.645
-1
0
1
2
3
4

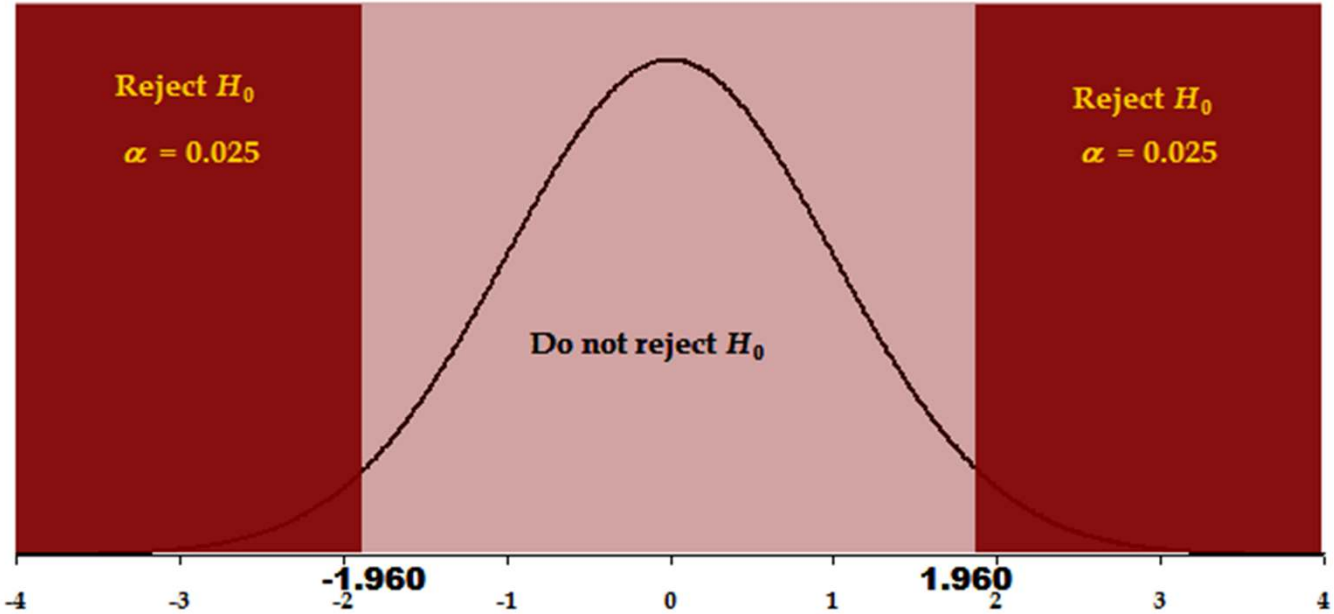

=NORMSINV(0.01)
C
D
E
-2.32635

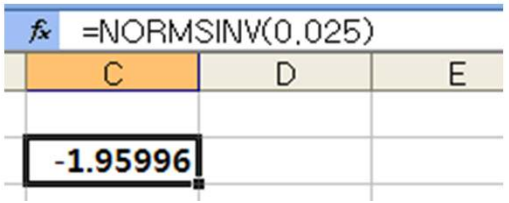

=NORMSINV(0.05)
C
D
E
-1.64485

Z-우측검정과 기각역


Reject H_0
α = 0.05
Reject H_0
α = 0.01
Do not reject H_0
-4
-3
-2
-1
0
1
1.645
2
2.326
3
4

Z-양측검정과 기각역


Reject H_0
α = 0.025
Reject H_0
α = 0.025
Do not reject H_0
-1.960
1.960
-4
-3
-2
-1
0
1
2
3
4


=NORMSINV(0.025)
C
D
E
-1.95996

문제

◆ **모 대학에서 통계학을 가르치는 황교수는 A가 20%, B가 30%, C가 40%, D 또는 F가 10%로 구성되도록 성적을 부여하고 있다. 총점의 평균이 60점 이고 표준편차가 10점인 정규분포를 따른다고 할 때**

- A를 받기 위해서는 총점이 몇 점 이상이어야 하는가?
- C를 받은 학생은 총점이 몇 점부터 몇 점 사이의 점수를 받았겠는가?

◆ **Hint : Excel의 normsinv 함수와 표준화 공식을 이용하세요.**