

6 표본분포

Topics:

- 6.1 서론
- 6.2 랜덤추출법
- 6.3 표본분포
- 6.4 표본분포의 모양과 중심극한정리

6.1 서론

Topics:

- 전수조사와 표본조사
- 확률표본추출과 확률표본

전수조사(census)와 표본조사(sample survey):

전수조사: _____의 모든 구성원을 대상으로 조사하여 모집단의 특성을 _____하는 방법
표본조사: _____을 추출하여 얻은 _____을 가지고 모집단의 특성을 _____하는 방법

- 예: 통계청에서 5년마다 인구주택총조사와 경제활동인구조사
- 예: 대선의 출구조사
- 표집분포(sampling distribution)란?
- 표집오차(sampling error)와 비표집오차(nonsampling error):

확률표본추출(probability sampling)과 확률표본:

모집단의 각 원소가 뽑힐 확률을 _____ 하고, 그 값들이 모두 _____ 보다 크게 되는 표본추출법

- 비확률표본추출이란?
- 예: 방송국 앞을 지나는 사람들을 표본으로 한 방송국의 여론조사는?
- 확률표본추출의 장점은?

6.2 랜덤추출법

Topics:

- 랜덤포본
- 난수표 사용
- 기타 표본추출방법

랜덤추출법(sampling)과 랜덤포본(random sample):

랜덤추출법: 모집단의 모든 원소가 표본으로 뽑힐 확률이 _____ 표본을 추출하는 방법
 랜덤포본: 랜덤추출법에 의해 추출된 표본

- 가능한 표본의 수와 특정 표본이 뽑힐 확률은?
- 예: 선거의 유권자 모집단에서 전화번호부를 사용해서 표본을 뽑으면?
- 예: 모집단 $\{a, b, c\}$ 에서 크기가 2인 표본을 뽑을 때 가능한 표본의 종류와 그 표본이 실제 표본으로 뽑힐 확률은?

	A	B	C	D	E	F	G	H	I	J	K
1	비복원추출										
2	가능한 표본										a, b, c가 실제 표본이 될 확률
3	표본으로 뽑힐 확률										
4											
5	복원추출										
6	가능한 표본										a, b, c가 한 번만 실제 표본이 될 확률
7	표본으로 뽑힐 확률										

난수표(random number table) 사용:

0 ~ 9까지의 숫자들이 _____의 상대도수를 가지면서 랜덤하게 나열되어 있는 난수들의 모임

• [표1] 난수표 참고

표 1 | 난수표

00000	10097	32533	76520	13586	34673	54876	80959	09117	39292	74945
00001	37542	04805	64894	74296	24805	24037	20636	10402	00822	91665
00002	08422	68953	19645	09303	23209	02560	15953	34764	35080	33606
00003	99019	02529	09376	70715	38311	31165	88676	74397	04436	27659
00004	12807	99970	80157	36147	64032	36653	98951	16877	12171	76833
00005	66065	74717	34072	76850	36697	36170	65813	39885	11199	29170
00006	31060	10805	45571	82406	35303	42614	86799	07439	23403	09732
00007	85269	77602	02051	65692	68665	74818	73053	85247	18623	88579
00008	63573	32135	05325	47048	90553	57548	28468	28709	83491	25624
00009	73796	45753	03529	64778	35808	34282	60935	20344	35273	88435
00010	98520	17767	14905	68607	22109	40558	60970	93433	50500	73998
00011	11805	05431	39808	27732	50725	68248	29405	24201	52775	67851
00012	83452	99634	06288	98083	13746	70078	18475	40610	68711	77817
00013	88685	40200	86507	58401	36766	67951	90364	76493	29609	11062
00014	99594	67348	87517	64969	91826	08928	93785	61368	23478	34113
00015	65481	17674	17468	50950	58047	76974	73039	57186	40218	16544
00016	80124	35635	17727	08015	45318	22374	21115	78253	14385	53763
00017	74350	99817	77402	77214	43236	00210	45521	64237	96286	02655
00018	69916	26803	66252	29148	36936	87203	76621	13990	94400	56418
00019	09893	20505	14225	68514	46427	56788	96297	78822	54382	14598
00020	91499	14523	68479	27686	46162	83554	94750	89923	37089	20048
00021	80336	94598	26940	36858	70297	34135	53140	33340	42050	82341
00022	44104	81949	85157	47954	32979	26575	57600	40881	22222	06413
00023	12550	73742	11100	02040	12860	74697	96644	89439	28707	25815
00024	63606	49329	16505	34484	40219	52563	43651	77082	07207	31790
00025	61196	90446	26457	47774	51924	33729	65394	59593	42582	60527
00026	15474	45266	95270	79953	59367	83848	82396	10118	33211	59466
00027	94557	28573	67897	54387	54622	44431	91190	42592	92927	45973
00028	42481	16213	97344	08721	16868	48767	03071	12059	25701	46670
00029	23523	78317	73208	89837	68935	91416	26252	29663	05522	82562
00030	04493	52494	75246	33824	45862	51025	61962	79335	65337	12472
00031	00549	97654	64051	88159	96119	63896	54692	82391	23287	29529
00032	35963	15307	26898	09554	33351	35462	77974	50024	90103	39333
00033	59808	08391	45427	26842	83609	49700	13021	24892	78565	20106
00034	46058	85236	01390	92286	77281	44077	93910	83647	70617	42941
00035	32179	00597	87379	25241	05567	07007	96743	17157	85394	11838
00036	69234	61406	20117	45204	15956	60000	18743	92423	97118	96338
00037	19565	41430	01758	75379	40419	21585	66674	36806	84962	85207
00038	45155	14938	19476	07246	43667	94543	59047	90033	20826	69541
00039	94864	31994	36168	10851	34888	81553	01540	35456	05014	51176
00040	98086	24826	45240	28404	44999	08896	39094	73407	35441	31880
00041	33185	16232	41941	50949	89435	48581	88695	41994	37548	73043
00042	80951	00406	96382	70774	20151	23387	96382	70774	25016	25298
00043	79752	49140	71961	28296	69861	02591	74852	20539	00387	59579
00044	18633	32537	98145	06571	31010	24674	05455	61427	77938	91936
00045	74029	43902	77557	32270	97790	17119	52527	58021	80814	51748
00046	54178	45611	80993	37143	05335	12969	36040	90324	36040	90324
00047	11664	49883	52079	84827	59381	71539	09973	33440	88461	23356
00048	48324	77928	31249	64710	02295	36870	32307	57546	15020	09994
00049	69074	94138	87637	91976	35584	04401	10518	21615	01848	76938

표 1 | 난수표

00000	10097	32533	76520	13586	34673	54876	80959	09117	39292	74945
00001	37542	04805	64894	74296	24805	24037	20636	10402	00822	91665
00002	08422	68953	19645	09303	23209	02560	15953	34764	35080	33606
00003	99019	02529	09376	70715	38311	31165	88676	74397	04436	27659
00004	12807	99970	80157	36147	64032	36653	98951	16877	12171	76833
00005	66065	74717	34072	76850	36697	36170	65813	39885	11199	29170
00006	31060	10805	45571	82406	35303	42614	86799	07439	23403	09732
00007	85269	77602	02051	65692	68665	74818	73053	85247	18623	88579
00008	63573	32135	05325	47048	90553	57548	28468	28709	83491	25624
00009	73796	45753	03529	64778	35808	34282	60935	20344	35273	88435
00010	98520	17767	14905	68607	22109	40558	60970	93433	50500	73998
00011	11805	05431	39808	27732	50725	68248	29405	24201	52775	67851
00012	83452	99634	06288	98083	13746	70078	18475	40610	68711	77817
00013	88685	40200	86507	58401	36766	67951	90364	76493	29609	11062
00014	99594	67348	87517	64969	91826	08928	93785	61368	23478	34113
00015	65481	17674	17468	50950	58047	76974	73039	57186	40218	16544
00016	80124	35635	17727	08015	45318	22374	21115	78253	14385	53763
00017	74350	99817	77402	77214	43236	00210	45521	64237	96286	02655
00018	69916	26803	66252	29148	36936	87203	76621	13990	94400	56418
00019	09893	20505	14225	68514	46427	56788	96297	78822	54382	14598
00020	91499	14523	68479	27686	46162	83554	94750	89923	37089	20048
00021	80336	94598	26940	36858	70297	34135	53140	33340	42050	82341
00022	44104	81949	85157	47954	32979	26575	57600	40881	22222	06413
00023	12550	73742	11100	02040	12860	74697	96644	89439	28707	25815
00024	63606	49329	16505	34484	40219	52563	43651	77082	07207	31790
00025	61196	90446	26457	47774	51924	33729	65394	59593	42582	60527
00026	15474	45266	95270	79953	59367	83848	82396	10118	33211	59466
00027	94557	28573	67897	54387	54622	44431	91190	42592	92927	45973
00028	42481	16213	97344	08721	16868	48767	03071	12059	25701	46670
00029	23523	78317	73208	89837	68935	91416	26252	29663	05522	82562
00030	04493	52494	75246	33824	45862	51025	61962	79335	65337	12472
00031	00549	97654	64051	88159	96119	63896	54692	82391	23287	29529
00032	35963	15307	26898	09554	33351	35462	77974	50024	90103	39333
00033	59808	08391	45427	26842	83609	49700	13021	24892	78565	20106
00034	46058	85236	01390	92286	77281	44077	93910	83647	70617	42941
00035	32179	00597	87379	25241	05567	07007	96743	17157	85394	11838
00036	69234	61406	20117	45204	15956	60000	18743	92423	97118	96338
00037	19565	41430	01758	75379	40419	21585	66674	36806	84962	85207
00038	45155	14938	19476	07246	43667	94543	59047	90033	20826	69541
00039	94864	31994	36168	10851	34888	81553	01540	35456	05014	51176
00040	98086	24826	45240	28404	44999	08896	39094	73407	35441	31880
00041	33185	16232	41941	50949	89435	48581	88695	41994	37548	73043
00042	80951	00406	96382	70774	20151	23387	96382	70774	25016	25298
00043	79752	49140	71961	28296	69861	02591	74852	20539	00387	59579
00044	18633	32537	98145	06571	31010	24674	05455	61427	77938	91936
00045	74029	43902	77557	32270	97790	17119	52527	58021	80814	51748
00046	54178	45611	80993	37143	05335	12969	36040	90324	36040	90324
00047	11664	49883	52079	84827	59381	71539	09973	33440	88461	23356
00048	48324	77928	31249	64710	02295	36870	32307	57546	15020	09994
00049	69074	94138	87637	91976	35584	04401	10518	21615	01848	76938

• 임의추출 방법의 예?

• 예: 균등분포를 따르는 모집단에서 표본크기가 2인 표본을 10번 추출하면?

	A	B	C	D	E	F	G	H	I	J	K
1	난수 생성	0.05626694	0.67840175	0.898234	0.61880529	0.26051723	0.51310663	0.78309056	0.40297367	0.77853041	0.76067783
2		0.71220616	0.04893196	0.39945238	0.59615218	0.52976225	0.71419309	0.44333739	0.17152055	0.74587407	0.90551761
3											
4											
5	평균	0.38423655	0.36366686	0.64884319	0.60747874	0.39513974	0.61364986	0.61321397	0.28724711	0.76220224	0.83309772

기타 표본추출방법:

단순랜덤추출(simple random sampling): 위의 무작위 추출

층화추출(stratified sampling): 모집단을 여러 개의 중복되지 않는 _____를 한 다음
각 층에서 _____ 표본을 산출하는 방법

집락추출(cluster sampling): 조사단위 대신 _____을 추출단위로 하는 추출방법

계통추출(systematic sampling): 추출간격(sampling interval) k 를 정해서
매 k 번째가 되는 단위들을 표본으로 선정하는 방법

6.3 표본분포

Topics:

- 표집분포
- 평균의 표집분포
- 표본평균의 기댓값
- 표본평균의 분산

표 집 분 포(sampling distribution):

한 모집단에서 같은 크기로 뽑은 모든 표본에서 통계량을 계산할 때, 이 통계량이 이루는 _____

- Recall.

평균의 표집분포:

관심 모수가 모집단의 평균이라면, 이를 _____으로 추정할 수 있다.

- 예: 1, 2, 3, 4라고 쓰여진 카드에서 2장을 뽑아서 그 평균을 확인하는 방법을 고려하자.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1																		
2	비복원추출																	
3	가능한 표본																	1, 2, 3, 4가 실제 표본이 될 확률
4	표본으로 뽑힐 확률																	
5	표본평균																	
6																		
7																		
8	복원추출																	
9	가능한 표본																	1, 2, 3, 4가 한 번만 실제 표본이 될 확률
10	표본으로 뽑힐 확률																	
11	표본평균																	
12																		

Y	Z	AA	AB
1. 모집단의 확률분포(N=4)			
x	f(x)	xf(x)	x ² f(x)
1			
2			
3			
4			
합계			
평균			
분산			

T	U	V	W
표본평균의 표집분포(비복원추출)			
\bar{x}	$f(\bar{x})$	$\bar{x}f(\bar{x})$	$\bar{x}^2f(\bar{x})$
1.5			
2			
2.5			
3			
3.5			
합계			
평균			
분산			
표본평균의 표집분포(복원추출)			
\bar{x}	$f(\bar{x})$	$\bar{x}f(\bar{x})$	$\bar{x}^2f(\bar{x})$
1			
1.5			
2			
2.5			
3			
3.5			
4			
합계			
평균			
분산			

표본평균의 기댓값:

$$\text{표본평균의 기댓값: } E(\bar{X}) =$$

- 모평균과 표본평균 사이의 관계: 모평균이 μ 인 임의의 모집단에서 크기 n 인 랜덤포본을 뽑을 때, 표본평균 \bar{X} 에 대하여 다음이 항상 성립한다.
- 예: 어느 대학교 남학생의 평균신장은 $170cm$ 로 알려져 있다. 이들 중에서 50명의 학생을 랜덤포본으로 뽑을 때, 표본평균은?
- 대수의 법칙(law of large number):

표본평균의 분산:

$$\text{Var}(\bar{X}) =$$

- 모분산과 표본평균의 분산과의 관계: 모분산이 σ^2 이고, 크기가 N 인 모집단에서 크기 n 인 랜덤포본을 뽑을 때 표본평균 \bar{X} 에 대하여 다음이 성립한다.

$$\text{비복원추출의 경우: } \text{Var}(\bar{X}) =$$

$$\text{복원추출의 경우: } \text{Var}(\bar{X}) =$$

- 예: 앞의 예제에 적용하면?

- 예: 어느 대학교 남학생 전체 5,000명의 평균신장은 170cm 이고, 분산은 9cm 으로 알려져 있다. 이들 중에서 50명의 학생을 랜덤포본으로 뽑을 때, 표본평균의 분산은?

- 표본오차(standard error):