

통계학 (강좌) 중간고사 2 (16:00~18:00)

※ 적절한 풀이과정이 없는 경우에는 정답으로 인정하지 않습니다.

※ 소수점 셋째자리까지 쓰세요.

1. (총 8점) 한 담배 제조회사에서는 새로 개발된 담배의 평균 타르 함량이 4mg 미만이라고 주장한다. 그 주장이 사실인가를 알아보기 위해 25개의 담배를 랜덤하게 선택해 타르 함량을 분석한 결과, 표본평균은 3.90mg, 표본표준편차는 0.14mg이었다. 다음 물음에 답하여라. 단, 전체 담배의 타르 함량은 정규분포를 따른다고 가정하자.

(1) (3점) 담배의 평균 타르 함량에 대한 95% 신뢰구간을 구하여라.

(2) (5점) 담배 회사의 주장은 타당하다고 할 수 있는가? 유의수준 5%에서 이를 검정하여라.

2. (8점) 토지 개발공사가 위촉한 두 명의 토지평가사 (A,B)의 감정가액에 차이가 있는가를 알아보기 위하여 8개의 특정지역을 선택하여 두 평가사에게 감정을 의뢰하여 구한 자료가 다음과 같다. 두 평가사의 감정결과가 동일하다고 볼 수 있는가에 대하여 유의수준 5%에서 검정을 실시하여라. 또한, 이 검정에 필요한 합리적인 가정을 함께 쓰시오.

지역	1	2	3	4	5	6	7	8
평가사 A	36	48	40	55	29	43	36	39
평가사 B	35	47	37	51	29	41	35	39

3. (총 17점) 제한속도가 100km/h인 고속도로에 과속을 방지하기 위해 무인 단속기를 설치하였다. 이 무인 단속기는 4개의 속도 측정 센서에서 측정한 속도 X_1, \dots, X_4 의 평균값

$$\bar{X} = \frac{1}{4} \sum_{i=1}^4 X_i$$

이 유의수준 5%에서 100km/h를 넘었다는 강한 증거가 있을 때 과속으로 판

정하고 있다. 자동차의 순간 속도가 μ km/h일 때 각각의 센서에서 측정되는 속도는 서로 독립이고 평균이 μ , 분산이 σ^2 인 정규분포를 따른다고 할 때 다음 물음에 답하시오.

(1) (2점) 자동차의 과속 여부를 판정하기 위한 귀무가설과 대립가설을 제시하시오.

(2) (5점) 모표준편차 $\sigma = 4$ (km/h)임이 알려져 있다고 하자. 이 때, 자동차를 과속으로 판정하게 되는 속도의 평균값의 범위를 구하시오.

(3) (5점) 이번에는 모표준편차를 모른다고 가정하자. 4개의 속도 측정 센서에서 측정한 속도의 표본 표준편차가 4(km/h) 일 때, 자동차를 과속으로 판정하게 되는 속도의 평균값의 범위를 구하시오.

(4) (5점) 실제 속도가 102.5km/h인 자동차를 과속이 아닌 것으로 판정하게 될 확률을 1% 이하로 줄이려면 몇 개의 센서를 사용해야 하는지 구하시오. 단, 모표준편차가 $\sigma = 4$ (km/h)임이 알려져 있다고 하자.

4. (총 13점) 모 마트에서는 개업 5주년을 맞이하여 추첨을 통해서 사은품을 지급하는 이벤트를 진행하려고 한다. 이를 위해 총 28개의 제비 중 당첨제비가 13개 들어 있는 주머니(A)를 사용한다고 공지를 하였다. 하지만 당첨 확률을 낮추기 위해 28개의 제비 중 당첨제비가 7개 들어 있는 주머니(B)로 바꿔치기를 했다는 제보가 있어 검사해 보려고 한다. 즉, 다음과 같은 가설을 검증하려고 한다.

$$H_0 : \text{주머니가 A이다} \quad \text{vs.} \quad H_1 : \text{주머니가 B이다.}$$

갑과 을은 각각 다음과 같은 방식으로 귀무가설(H_0)의 기각여부를 결정하자고 제안했다.

갑 : “주머니에서 제비를 임의로 2개 뽑아, 2개 모두 비당첨제비이면 귀무가설을 기각하겠다.”
 을 : “두 개의 주사위를 동시에 던져서 눈의 합이 5 이하이면 귀무가설을 기각하겠다.”

- (1) (5점) 갑, 을의 방식으로 가설을 검증할 때, 제 1종의 오류를 범할 확률을 각각 구하라.
- (2) (5점) 갑, 을의 방식으로 가설을 검증할 때, 제 2종의 오류를 범할 확률을 각각 구하라.
- (3) (3점) 위 결과를 바탕으로 갑과 을 중에 누가 제안한 검증 방식이 더 좋은지를 간략하게 기술하시오.

5. (8점) 연령이 비슷한 정상인(x_1) 10명과 정신장애인(x_2) 10명에 대하여 뇌세포 조직을 조사하여 특정효소활동에 의해 1시간 동안 생성되는 물질의 양을 조사한 결과 다음의 결과를 얻었다.

$$\begin{array}{ll} \bar{x}_1 = 42.3 & s_1 = 12.1 \\ \bar{x}_2 = 36.5 & s_2 = 5.1 \end{array}$$

정상인의 평균수치가 정신장애인의 평균수치보다 더 크다고 할 수 있는지 유의수준 5%에서 검정하시오. 두 모집단의 분포는 정규분포를 따른다고 가정하자.

또한, 필요시 다음과 같은 t 분포의 근사자유도를 사용하시오.

$$df = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{1}{n_1-1}(S_1^2/n_1)^2 + \frac{1}{n_2-1}(S_2^2/n_2)^2}$$

6. (총 8점) 최근에 한 조사기관에서는 대학졸업자 중에서 20%가 자신의 전공을 살릴 수 있는 직장에 입사한다고 발표하였다. 그러나 대학졸업자들이 결과에 대해 이의를 제기하자 이 발표가 사실인가를 알아보기 위하여 400명의 대학졸업자를 임의로 선발하여 조사한 결과 100명이 전공을 살릴 수 있는 직장에서 일하고 있었다.

- (1) (5점) 조사기관의 발표가 타당하다고 할 수 있는지를 유의수준 5%에서 검정하여라.
- (2) (3점) 모비율에 대한 95% 신뢰구간을 구하시오.

7. (총 20점) A과목에 대한 새로운 학습법이 개발되었다. 새롭게 개발된 학습법을 통한 학생들의 학습 성취도를 평가하려 한다. 정확한 평가를 위해 O/X 문제를 출제하여 A과목에 대한 사전지식이 있는 학생들을 실험 대상에서 제외하려고 한다. “사전 지식이 전혀 없는 학생은 O/X문제의 정답률이 50%이다” 라고 가정하고 다음 물음에 답하시오.

(1) (5점) (표본의 크기가 작은 경우) 학생들에게 O/X 문제를 5문제씩 출제하여 A과목에 대한 사전지식이 있는다는 강한 증거가 있을 때 해당 학생을 실험 대상에서 제외한다고 할 때 몇 문제 이상 맞힌 학생을 제외하는 것이 타당한지 설명하시오. (단, 유의수준은 5%)

(2) (5점) (1)과 같은 방법으로 사전지식이 있는 학생을 제외하였을 때 실제로는 사전지식이 있어 정답률이 80%인 학생을 실험대상에서 제외하지 못하게 될 확률을 구하시오.

(3) (5점) (표본의 크기가 큰 경우) (2)와 같은 문제점을 파악한 실험자는 문제수를 25문제로 늘렸다고 한다. 이제 몇 문제 이상 맞힌 학생을 실험대상에서 제외하는 것이 좋은지 구하고, 이 경우 실제로는 사전지식이 있어 정답률이 80%인 학생을 실험대상에서 제외하지 못하게 될 확률을 구하시오.

(4) (5점) (3)과 같은 방법으로 선택한 실험대상 학생 5명의 점수가 다음과 같다. 이 학생들의 정답률에 차이가 있다고 할 수 있는지 검정하시오. (단, 유의수준은 5%)

	학생1	학생2	학생3	학생4	학생5	평균
정답 수	12	16	11	15	11	13
오답 수	13	9	14	10	14	12

8. (8점) 어느 회사에서 1일 3교대로 작업을 하고 있다. 작업조에 따라 불량률이 다른지를 알아보기 위해 주간조에서 생산한 150개, 야간조에서 120개와 심야조에서 130개를 뽑아 불량품과 양호품의 개수를 조사한 결과가 다음과 같다.

작업조 상태	주간	야간	심야	합계
불량품	25	20	15	60
양호품	125	100	115	340
합계	150	120	130	400

작업조에 따른 제품의 불량률은 차이가 있는가? 유의수준 5%에서 이를 검정하시오.

9. (총 10점, 각 2점) 다음 명제에 대하여 맞으면 O, 틀리면 X로 답하시오.

(1) 유의수준 0.05에서 귀무가설 H_0 을 기각시키지 못했더라도 유의수준 0.01에서 귀무가설 H_0 을 기각시킬 수 있다.()

(2) 유의확률이 0.1일 때 귀무가설을 기각하게 되는 유의수준의 최대값은 0.1이다.()

(3) 가설검정에서 제 1종의 오류를 범할 확률과 제 2종의 오류를 범할 확률의 합은 1이다.()

(4) 모비율에 대한 이표본 z-검정과 카이제곱통계량을 이용한 2×2 분할표에 대한 동질성 검정의 결과는 일치한다. ()

(5) 표본분산과 모분산이 정확히 일치하는 경우 t-분포를 이용해 구한 신뢰구간이 정규분포를 이용하여 구한 신뢰구간에 비해 더 넓다. ()

표준 정규 분포표 $P(Z \leq z)$ [illegible]

t 분포표

$t_\alpha : P(T \geq t_\alpha) = \alpha, T \sim t(df)$

df \ α	0.10	0.05	0.025	0.01
1	3.078	6.31	12.71	31.82
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602
16	1.337	1.746	2.120	2.583
17	1.333	1.740	2.110	2.567
18	1.330	1.734	2.101	2.552
19	1.328	1.729	2.093	2.539
20	1.325	1.725	2.086	2.528
21	1.323	1.721	2.080	2.518
22	1.321	1.717	2.074	2.508
23	1.319	1.714	2.069	2.500
24	1.318	1.711	2.064	2.492
25	1.316	1.708	2.060	2.485

χ^2 분포표

$\chi_\alpha^2 : P(\chi^2 \geq \chi_\alpha^2) = \alpha, \chi^2 \sim \chi^2(df)$

df \ α	0.975	0.95	0.90	0.10	0.05	0.025
1	0.00	0.00	0.02	2.71	3.84	5.02
2	0.05	0.10	0.21	4.61	5.99	7.38
3	0.22	0.35	0.58	6.25	7.81	9.35
4	0.48	0.71	1.06	7.78	9.49	11.14
5	0.83	1.15	1.61	9.24	11.07	12.83
6	1.24	1.64	2.20	10.64	12.59	14.45
7	1.69	2.17	2.83	12.02	14.07	16.01
8	2.18	2.73	3.49	13.36	15.51	17.53
9	2.70	3.33	4.17	14.68	16.92	19.02
10	3.25	3.94	4.87	15.99	18.31	20.48
11	3.82	4.57	5.58	17.28	19.68	21.92
12	4.40	5.23	6.30	18.55	21.03	23.34
13	5.01	5.89	7.04	19.81	22.36	24.74
14	5.63	6.57	7.79	21.06	23.68	26.12
15	6.26	7.26	8.55	22.31	25.00	27.49
16	6.91	7.96	9.31	23.54	26.30	28.85
17	7.56	8.67	10.09	24.77	27.59	30.19
18	8.23	9.39	10.86	25.99	28.87	31.53
19	8.91	10.12	11.65	27.02	30.14	32.85
20	9.59	10.85	12.44	28.41	31.41	34.17
21	10.28	11.59	13.24	29.62	32.67	35.48
22	10.98	12.34	14.04	30.81	33.92	36.78
23	11.69	13.09	14.85	32.01	35.17	38.08
24	12.40	13.85	15.66	33.20	36.42	39.36
25	13.12	14.61	16.47	34.38	37.65	40.65

F 분포표

$$F_{0.05} : P(F \geq F_{0.05}) = 0.05, F \sim F(df_1, df_2)$$

	df_1												
df_2	1	2	3	4	5	6	7	8	9	10	11	12	13
1	161.4	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91	244.69
2	18.51	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396	19.405	19.413	19.419
3	10.12	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786	8.763	8.745	8.729
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.936	5.912	5.891
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.704	4.678	4.655
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	4.027	4.000	3.976
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.603	3.575	3.550
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347	3.313	3.284	3.259
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.102	3.073	3.048
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.943	2.913	2.887
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.818	2.788	2.761
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.717	2.687	2.660
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.635	2.604	2.577
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.565	2.534	2.507
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.507	2.475	2.448

F 분포표

$$F_{0.025} : P(F \geq F_{0.025}) = 0.025, F \sim F(df_1, df_2)$$

	df_1												
df_2	1	2	3	4	5	6	7	8	9	10	11	12	13
1	647.7	799.50	864.16	899.58	921.84	937.11	948.21	956.65	963.28	968.62	973.02	976.70	979.83
2	38.50	39.000	39.165	39.248	39.298	39.331	39.355	39.373	39.387	39.398	39.407	39.415	39.421
3	17.44	16.044	15.439	15.101	14.885	14.735	14.624	14.540	14.473	14.419	14.374	14.337	14.304
4	12.21	10.649	9.979	9.605	9.364	9.197	9.074	8.980	8.905	8.844	8.794	8.751	8.715
5	10.00	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619	6.568	6.525	6.488
6	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523	5.461	5.410	5.366	5.329
7	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823	4.761	4.709	4.666	4.628
8	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295	4.243	4.200	4.162
9	7.209	5.715	5.078	4.718	4.484	4.320	4.197	4.102	4.026	3.964	3.912	3.868	3.831
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717	3.665	3.621	3.583
11	6.724	5.256	4.630	4.275	4.044	3.881	3.759	3.664	3.588	3.526	3.474	3.430	3.392
12	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436	3.374	3.321	3.277	3.239
13	6.414	4.965	4.347	3.996	3.767	3.604	3.483	3.388	3.312	3.250	3.197	3.153	3.115
14	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3.285	3.209	3.147	3.095	3.050	3.012
15	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060	3.008	2.963	2.925