

# 통 계 학

▶ 응시번호 :

▶ 성 명 :

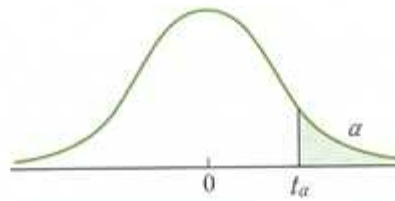
## 제 1 문 (15점)

어느 건강 클리닉에서 다이어트 프로그램 A와 B의 체중 감량 효과를 비교하기 위하여 지원자 18명을 임의로 각 프로그램에 9명씩 할당하고 6개월간 각 다이어트 프로그램을 수행한 후 체중 감량분을 조사하여 다음과 같은 결과를 얻었다.

	표본평균	표본분산
프로그램 A	14	19
프로그램 B	11	17

- (1) 프로그램 A와 프로그램 B의 합동분산(pooled variance)의 추정치를 구하라.
- (2) 프로그램 A에서의 체중감소분의 분산과 프로그램 B에서의 체중감소분의 분산이 같다고 가정할 때  
 $H_0$ : 두 다이어트 프로그램에 대한 체중감량 평균의 차이가 없다  
 $H_1$ : 두 다이어트 프로그램에 대한 체중감량 평균의 차이가 있다  
 에 대한 검정을 유의수준 5%에서 실시하라.
- (3) 위의 자료에 대하여 성별, 프로그램 시작 전 체중, 체형 등을 고려하여 프로그램 A의 지원자와 프로그램 B의 지원자를 짝을 지운 후 표본상관계수를 계산하였더니 0.8이었다. 각 짝에서 프로그램 A 수강생의 감량분에서 프로그램 B의 감량분을 뺀 값에 대한 평균은 3, 표본분산은 7.24, 표본표준편차는 2.69이었다. 이 경우에 대하여 (2)에서와 같은 가설을 유의수준 5%에서 검정하라.
- (4) (2)의 결과와 (3)의 결과를 비교하여 논하라.

<표> t분포의 상위  $\alpha$ 의 확률을 주는 값



d.f. \ $\alpha$	.25	.10	.05	.025	.01	.00833	.00625	.005
1	1.000	3.078	6.314	12.706	31.821	38.190	50.923	63.657
2	.816	1.886	2.920	4.303	6.965	7.649	8.860	9.925
3	.765	1.638	2.353	3.182	4.541	4.857	5.392	5.841
4	.741	1.533	2.132	2.776	3.747	3.961	4.315	4.604
5	.727	1.476	2.015	2.571	3.365	3.534	3.810	4.032
6	.718	1.440	1.943	2.447	3.143	3.287	3.521	3.707
7	.711	1.415	1.895	2.365	2.998	3.128	3.335	3.499
8	.706	1.397	1.860	2.306	2.896	3.016	3.206	3.355
9	.703	1.383	1.833	2.262	2.821	2.933	3.111	3.250
10	.700	1.372	1.812	2.228	2.764	2.870	3.038	3.169
11	.697	1.363	1.796	2.201	2.718	2.820	2.981	3.106
12	.695	1.356	1.782	2.179	2.681	2.779	2.934	3.055
13	.694	1.350	1.771	2.160	2.650	2.746	2.896	3.012
14	.692	1.345	1.761	2.145	2.624	2.718	2.864	2.977
15	.691	1.341	1.753	2.131	2.602	2.694	2.837	2.947
16	.690	1.337	1.746	2.120	2.583	2.673	2.813	2.921
17	.689	1.333	1.740	2.110	2.567	2.655	2.793	2.898
18	.688	1.330	1.734	2.101	2.552	2.639	2.775	2.878
19	.688	1.328	1.729	2.093	2.539	2.625	2.759	2.861
20	.687	1.325	1.725	2.086	2.528	2.613	2.744	2.845
21	.686	1.323	1.721	2.080	2.518	2.601	2.732	2.831
22	.686	1.321	1.717	2.074	2.508	2.591	2.720	2.819
23	.685	1.319	1.714	2.069	2.500	2.582	2.710	2.807
24	.685	1.318	1.711	2.064	2.492	2.574	2.700	2.797
25	.684	1.316	1.708	2.060	2.485	2.566	2.692	2.787
26	.684	1.315	1.706	2.056	2.479	2.559	2.684	2.779
27	.684	1.314	1.703	2.052	2.473	2.552	2.676	2.771
28	.683	1.313	1.701	2.048	2.467	2.546	2.669	2.763
29	.683	1.311	1.699	2.045	2.462	2.541	2.663	2.756
30	.683	1.310	1.697	2.042	2.457	2.536	2.657	2.750
40	.681	1.303	1.684	2.021	2.423	2.499	2.616	2.704
60	.679	1.296	1.671	2.000	2.390	2.463	2.575	2.660
120	.677	1.289	1.658	1.980	2.358	2.428	2.536	2.617
$\infty$	.674	1.282	1.645	1.960	2.326	2.394	2.498	2.576

## 제 2 문 (20점)

어느 지역에서 소득에 대하여 알아보기 위하여 소득( $Y$ ), 연령( $X_1$ ), 교육기간( $X_2$ )을 조사하였다. 조사한 자료에서  $Y$ 를 반응변수(종속변수),  $X_1$ ,  $X_2$ 를 설명변수(독립변수)로 하여 회귀분석을 실시한 결과가 다음과 같다. 각 회귀모형에서  $\epsilon_i \sim iid N(0, \sigma^2)$ 라고 가정한다.

모형 1 :  $y_i = \beta_0 + \beta_1 x_1 + \epsilon_i$

Source	자유도	제곱합	평균제곱합	F-값
$X_1$	1	19.9558	19.95586	20.79
오 차	98	94.0783	0.95998	
전 체	99	114.0343		
	Estimate	Std. Error	t value	Pr >  t
Intercept	1.86987	0.40242	4.65	<.0001
$X_1$	0.37845	0.08301	4.56	<.0001

모형 2 :  $y_i = \beta_0 + \beta_2 x_2 + \epsilon_i$

Source	자유도	제곱합	평균제곱합	F-값
$X_2$	1	24.0222	24.0222	26.15
오 차	98	90.0121	0.9185	
전 체	99	114.0343		
	Estimate	Std. Error	t value	Pr >  t
Intercept	0.8294	0.16217	5.11	<.0001
$X_2$	0.6505	0.04624	14.07	<.0001

모형 3 :  $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon_i$

Source	자유도	제곱합	평균제곱합	F-값
$X_1, X_2$	2	112.8308	56.41539	4547.07
오 차	97	1.2035	0.01241	
전 체	99	114.0343		

	Estimate	Std. Error	t value	Pr >  t	VIF
Intercept	3.8133	0.05097	74.82	<.0001	0
$X_1$	-1.0114	0.01863	-54.29	<.0001	3.89775
$X_2$	1.4172	0.01638	86.52	<.0001	3.89775

- (1)  $Y$ 의 변동에 대하여  $X_2$ 가 모형에 있을 때  $X_1$ 이 추가됨으로써 설명이 증가하는 부분은 얼마인가?
- (2) 모형 3에서 독립변수(또는 설명변수)  $X_1$ 이 종속변수(또는 반응변수)  $Y$ 에 미치는 영향이 유의한지를 검정하고  $X_1$ 이  $Y$ 에 미치는 영향을 설명하라.
- (3) 모형 1에서의  $\beta_1$  추정값과 모형 3에서의  $\beta_1$  추정값의 부호가 다른 이유를 설명하라.
- (4) 추후 모형 3에서 성별을 추가하여 소득에 미치는 영향을 보려고 한다. 이를 위한 모형을 설정하고 모형을 설명하라.

---

### 제 3 문 (10점)

---

은행 창구에서 고객이 기다리는 시간을  $X$ 라고 하면 확률변수  $X$ 는 지수분포를 따르며 확률분포 함수  $f(x)$ 는  $f(x) = ce^{-cx}$ ,  $x \geq 0$ 이다. 한 라인에서의 고객 평균 대기시간은 4분으로 알려져 있다. 다음에 답하라.

- (1) 고객이 기다리는 시간이 6분을 넘어갈 확률은 구하라.
- (2)  $k$ 개( $k \geq 1$ )의 창구를 열어 적어도 한 고객이 4분 이상 기다리지 않을 확률이 0.95 이상이 되도록 하고자 한다. 이 때 필요한 최소의 값  $k$ 를 구하라.

$x$	-2.0	-1.9	-1.8	-1.7	-1.6	-1.5	-1.4	-1.3	-1.2	-1.1	-1.0
$e^x$	0.14	0.15	0.17	0.18	0.20	0.22	0.25	0.27	0.30	0.33	0.37

---

### 제 4 문 (5점)

---

두 개의 동전을 차례로 던진다. 좌표평면 위의 점  $A(x, y)$ 는 원점에서 출발하여  $x$ 는 첫 번째 동전이 앞면이면 +1, 뒷면이면 -1만큼 움직이고,  $y$ 는 두 번째 동전이 앞면이면 +1, 뒷면이면 -1만큼 움직인다. 두 개의 동전을 던지는 시행을 10번 독립적으로 반복 시행했을 때 점  $A$ 가 중심이 원점이고 반지름이 1인 원 위 또는 원 이내에 있을 확률을 구하라.