FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS INDONESIA

MIDTERM EXAM 2017/2018
ADVANCE STATISTICS - SPECIAL CLASS

Date : Monday, October l6th 2017
Time : 150 minutes
Lecturer : Dewi Ratna Sjari M.

You can use ordinary calculator - Its is strictly prohibited to use cellular phone as calculator!!

A. Theoritical Problems (15 points):

1. The chi-square test of a contingency table is a test of independence for:
   A. A single qualitative variable
   B. Two qualitative variables
   C. Two quantitative variables
   D. Three or more quantitative variables

2. If there are five treatments under study, the number of pairwise comparisons is:
   A. 15 B. 5
   C. 20 D. 10

3. The chi-square distribution can assume
   A. only positive values.
   B. only negative values.
   C. negative and positive values or zero.
   D. only zero.

4. Two chi-square distributions were plotted on the same chart. One distribution was for 3 degrees of freedom and the other was for 12 degrees of freedom. Which distribution would tend to approach a normal distribution?
   A. 3 degrees B. 12 degrees
   C. 9 degrees D. All would

5. In a two-way ANOVA, a blocking variable is used to
   A. increase the error sum of squares.
   B. decrease the error sum of squares.
   C. increase the treatment sum of squares.
   D. decrease the treatment sum of squares.

6. If the decision is to reject the null hypothesis of no difference between two population proportions at the 10% level of significance, what are the alternate hypothesis and rejection region?
   A. $\pi_1 \neq \pi_2; z > 1.65$ and $z < -1.65$
   B. $\pi_1 \neq \pi_2; z > 1.96$ and $z < -1.96$
   C. $\pi_1 > \pi_2; < -1.65$
   D. $\pi_1 > \pi_2; < -1.96$

7. For the chi-square test of a contingency table, the expected cell frequencies are found as $e_{ij} = \frac{(Row\ i\ Total)(Column\ j\ Total)}{sample\ Size}$ which is the same as:

A. The observed cell frequencies
B. The cell probability multiplied by the sample size
C. The row total
D. The column total

8. Psychology students want to determine if there are differences between the ability of business majors and science majors to solve various types of analytic problems. They conduct an experiment and record the amount of time it takes to complete each analytic problem. The following two-way ANOVA table summarizes their findings.

Source of Variation	*SS*	*df*	*MS*	*F*
Major	16.82	1	16.82	11.21
Problem	102.92	4	25.73	17.15
Interaction	157.48	4	39.37	26.25
Error	60.00	40	1.50	
Total	337.22	49		

The number of different analytic problems the students solved is:
A. 1 B. 2
C. 4 D. 5

9. The variability due to chance, also known as within-treatments variability, is the estimate of σ^2 which is based on the:
A. variability of the data across different samples
B. consistency of the data within each sample
C. variability of the data within each sample
D. reliability of the data within each sample

10. If the amount of variability between treatments is significantly greater than the amount of variability within treatments, then:
A. reject the null hypothesis of equal population means
B. do not reject the null hypothesis of equal population means
C. conclude that the ratio of between-treatments variability to within-treatments variability is significantly less than 1
D. perform further analysis using the two-way ANOVA with interaction

11. In a market test of a new chocolate raspberry coffee, a poll of 400 people from Dobbs Ferry showed 250 preferred the new coffee. In Irvington, 170 out of 350 people preferred the new coffee. To test the hypothesis that there is no difference in preferences between the two villages, what is the alternate hypothesis?
A. $\pi_1 < \pi_2$ B. $\pi_1 > \pi_2$
C. $\pi_1 = \pi_2$ D. $\pi_1 \neq \pi_2$

12. The chi-square test of a contingency table is valid when the expected cell frequencies are:
A. Equal to 0 B. More than 0 but less than 5
C. At least 5 D. Negative

13. $P\left(\chi^2_{1-w/2,v} < \chi^2_1 < \chi^2_{w/2,v}\right) = 0.95$ If $v = 7$, then the value of $\chi^2_{1-w/2,v}$ is: . . . . . .

14. $P\left(F_{1-w/2,v1,v2} < F_1 < F_{w/2,v1,v2}\right) = 0.90$ If $v_1 = 5$, and $v_2 = 4$, then the value of $F_{w/2,v1,v2}$ is: . . . . . .

15. $P\left(\chi^2_1 < \chi^2_{w,v}\right) = 0.95.$ If $v = 10$, then the value of $\chi^2_{w,v}$ is: . . . . . .

B. Problematic Questions:

1. **(20 points)** Karier.com conducts an annual salary survey of banking professionals. For the the 2013 (March) survey, 16 responses submitted online to Web-based survey by professionals in Jabaar city. 50 % of the respondents were women. The company as usual intends to construct salary comparissons by gender. Last year they found the dispersion of wage between male and female professional were quite similar. The data of respondent wage (in million rupiah) are shown in the following table:

Male	9	8	14	10	12	7	9	11
Female	7	5	9	8	6	7	7	5

   a. Is there any sufficient evidence of difference between mean salary of male and female banking professional, particularly in Jabaar city? Test using $\alpha = 0.1$. Interpret your result.
   b. What are the assumption necessary for this test?

2. **(20 points)** Recently Centre-One has done a survey of supermarket chains. Random samples of 150 employees were interviewed. There are 3 group of employees, those who work at the supermarket's warehouse (warehouse staff=WS), those who work as sales staff (SS) and group of workers who work in the office administrative (AS). The three groups were asked about their opinion concerning the new minimum salary policy, with a scale, namely: Favored; preferred; disliked; and bad. Of the 40 WS interviewed, 7 answered 'preferred', 24 answered 'dislike', and 8 responded 'bad'. Of 50 sales staff, 8 responded 'bad', 9 answered 'preferred', and 3 responded 'Favored'. The rest of the samples are administrative staff, of which 14 replied 'Favored' and 2 answered 'bad'. In overall survey, most of the employees interviewed responded negatively. With a confidence level of 95%. Could you help Centre-One to conclude if employee attitudes in response to the salary restrictions depending on the type / class of employee?

	Favored	**Preferred**	**Dislike**	**Bad**	**Total**
WS	1	7	24	8	40
SS	3	9	30	8	50
SA	14	20	24	2	60
Total	18	36	78	18	150

3. **(20 points)** An experiment to determine the most effective way to teach safety principles applied four different teaching methods. Some employees were given programmed instruction booklets and worked through the course at their own pace. Other employees attended lectures. A third group watched a television presentation, and a fourth group was divided into small discussion groups. A high of 10 was possible. A sample of five tests was selected from each group. The test grade results were:

Program Instruction	Lecture	TV	Group Discussion
6	8	7	8
7	5	9	5
6	8	6	6
5	6	8	6
6	8	5	5

   1. Prepare and complete the ANOVA table.
   2. At the 5% significance level, what is your conclusion for the hypothesis test about the methods?
   3. Explain whether the p-value for the test about brand is below or above 0.05 (use F test).

4. **(25 points)** The Indonesian for Women empowerment association recently conducted a study to compare the monthly wages of men and women employed in agriculture, industry, and service sector.

   Questions:
   a. Complete the table below!
   b. Conduct a test to confirmed the hyphothesis that gender wages are the same and the sector wages are the same
   c. Test the interaction efect of gender and sector on wages
   d. Interpret the result in a brief report

	Sector		
Gender	Agriculture	Manufacture	Service
Men	978	1030	1335
	1035	1095	1167
	964	1135	1075
	1010	1005	1218
	1117	1169	1065
Women	896	965	1079
	975	960	1000
	999	1038	1006
	1019	1130	1110
	1037	1020	1085

SUMMARY	Agriculture	Manufacture	Service	Total
Men				
Count	5	5	5	15
Sum	5104	5434	5860	16398
Average	1020,8	1086,8	1172	1093,2
Variance	3657,7	4766,2	12392	10051,74

Women				
Count	5	5	5	15
Sum	4926	5113	5280	15319
Average	985,2	1022,6	1056	1021,267
Variance	3019,2	4753,8	2480,5	3825,638

Total			
Count	10	10	10
Sum	10030	10547	11140
Average	1003	1054,7	1114
Variance	3319,556	5376,011	10347,78

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*F crit*
Sample	38808,03	.....	.....	.....	.....
Columns	.....	.....	.....	.....	.....
Interaction	.....	.....	4152,233	.....	.....
Within	.....	.....	.....	.....	
Total	.....				

Annexes

Selected Formulas:

1. $\chi^2_{Hit} = \sum\left(\frac{(f_{oi} - f_{ei})^2}{f_{ei}}\right)$

2. $SSTotal = \sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(X_{ij} - \bar{X}\right)^2$ $\quad SSTreatment = \sum_{i=1}^{k} n_i\left(\bar{X}_i - \bar{X}\right)^2$

3. $SSTotal = \sum_{i=1}^{K}\sum_{j=1}^{H}\left(X_{ij} - \bar{X}\right)^2$ $\quad SSGroup = H\sum_{i=1}^{K}\left(\bar{X}_{i\bullet} - \bar{X}\right)^2$

   $SSBlock = K\sum_{j=1}^{H}\left(\bar{X}_{\bullet j} - \bar{X}\right)^2$

   Mean dari Group *i*: $\bar{X}_{i\bullet} = \frac{\sum_{j=1}^{H} X_{ij}}{H}$ $\quad$ Mean dari Block *j*: $\bar{X}_{\bullet j} = \frac{\sum_{i=1}^{K} X_{ij}}{K}$

4. $SST = \sum_{i=1}^{K}\sum_{j=1}^{H}\sum_{l=1}^{L}\left(X_{ijl} - \bar{X}\right)^2$ $\quad SSG = HL\sum_{i=1}\left(\bar{X}_{i\bullet\bullet} - \bar{X}\right)^2$

   $SSI = L\sum_{i=1}\sum_{j=1}\left(\bar{X}_{ij\bullet} - \bar{X}_{i\bullet\bullet} - \bar{X}_{\bullet j\bullet} + \bar{X}\right)^2$ $\quad SSB = KL\sum_{j=1}\left(\bar{X}_{\bullet j\bullet} - \bar{X}\right)^2$

   Group Means : $\bar{X}_{i\bullet\bullet} = \frac{\sum_{j=1}^{H}\sum_{l=1}^{L} X_{ijl}}{HL}$ $\quad$ Block Means : $\bar{X}_{\bullet j\bullet} = \frac{\sum_{i=1}^{K}\sum_{l=1}^{L} X_{ijl}}{KL}$

   Cell Means : $\bar{X}_{ij\bullet} = \frac{\sum_{l=1}^{L} X_{ijl}}{L}$

5. $\sigma_{\bar{X}_1 - \bar{X}_2} = \sigma\sqrt{\frac{1}{n1} + \frac{1}{n2}}$

$$\sigma_{\bar{X}_1-\bar{X}_2} = s_p\sqrt{\frac{1}{n1}+\frac{1}{n2}} = \sqrt{s_p^2\left(\frac{1}{n1}+\frac{1}{n2}\right)}\ ;\ s_p^2 = \frac{(n1-1)s_{X1}^2+(n2-1)s_{X2}^2}{n1+n2-2}$$

$$\sigma_{p_1-p_2} = \sqrt{\frac{p_c(1-p_c)}{n1}+\frac{p_c(1-p_c)}{n2}};\ p_c = \frac{x_1+x_n}{n1+n2}$$

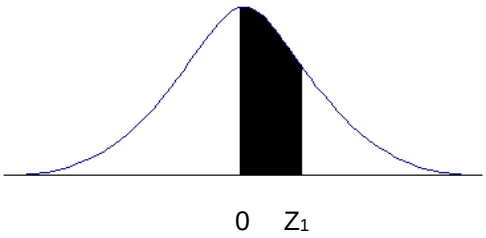
$$\sigma_{p_1-p_2} = \sqrt{\frac{p_1(1-p_1)}{n1}+\frac{p_2(1-p_2)}{n2}}$$

Kanopi FEBUI

Unity in Development

Standard Normal Distribution:

The table shows the area under the curve or probability of Z between 0 and $[P(0 < Z < Z_1)]$
Use symmetricity for negative values.
Value of Z_1 is shown by the total of values at the left and upper margins



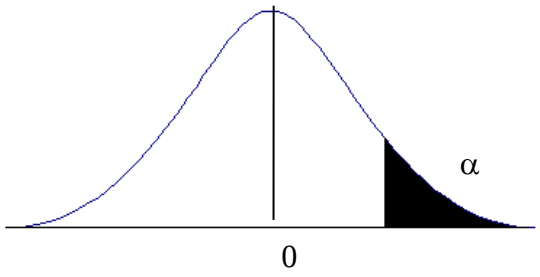
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Kanopi FEBUI

Unity in Development

Student *t* Distribution:

The table shows that at the degree of freedom at left margin then the probability of *t* will be greater than the content in the table is α as shown at upper margin [$P(t > t_1) = \alpha$]. Use symmetricity for negative values.
For df = 12 then P (t > 1.782) = 0.05; and
for df = 12 then P (t < – 1.782) = 0.05

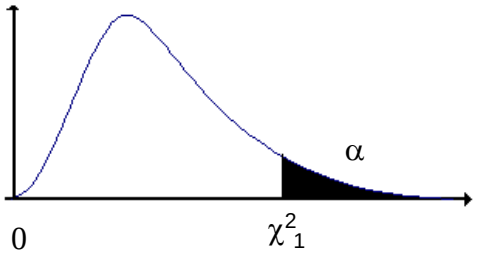


df \ α	0.1	0.05	0.025	0.01	0.005
1	3.0777	6.3137	12.7062	31.8210	63.6559
2	1.8856	2.9200	4.3027	6.9645	9.9250
3	1.6377	2.3534	3.1824	4.5407	5.8408
4	1.5332	2.1318	2.7765	3.7469	4.6041
5	1.4759	2.0150	2.5706	3.3649	4.0321
6	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.4149	1.8946	2.3646	2.9979	3.4995
8	1.3968	1.8595	2.3060	2.8965	3.3554
9	1.3830	1.8331	2.2622	2.8214	3.2498
10	1.3722	1.8125	2.2281	2.7638	3.1693
11	1.3634	1.7959	2.2010	2.7181	3.1058
12	1.3562	1.7823	2.1788	2.6810	3.0545
13	1.3502	1.7709	2.1604	2.6503	3.0123
14	1.3450	1.7613	2.1448	2.6245	2.9768
15	1.3406	1.7531	2.1315	2.6025	2.9467
16	1.3368	1.7459	2.1199	2.5835	2.9208
17	1.3334	1.7396	2.1098	2.5669	2.8982
18	1.3304	1.7341	2.1009	2.5524	2.8784
19	1.3277	1.7291	2.0930	2.5395	2.8609
20	1.3253	1.7247	2.0860	2.5280	2.8453
21	1.3232	1.7207	2.0796	2.5176	2.8314
22	1.3212	1.7171	2.0739	2.5083	2.8188
23	1.3195	1.7139	2.0687	2.4999	2.8073
24	1.3178	1.7109	2.0639	2.4922	2.7970
25	1.3163	1.7081	2.0595	2.4851	2.7874
26	1.3150	1.7056	2.0555	2.4786	2.7787
27	1.3137	1.7033	2.0518	2.4727	2.7707
28	1.3125	1.7011	2.0484	2.4671	2.7633
29	1.3114	1.6991	2.0452	2.4620	2.7564
30	1.3104	1.6973	2.0423	2.4573	2.7500
40	1.3031	1.6839	2.0211	2.4233	2.7045
50	1.2987	1.6759	2.0086	2.4033	2.6778
60	1.2958	1.6706	2.0003	2.3901	2.6603
80	1.2922	1.6641	1.9901	2.3739	2.6387
100	1.2901	1.6602	1.9840	2.3642	2.6259
120	1.2886	1.6576	1.9799	2.3578	2.6174

Chi Square (χ^2) Distribution:

The table shows that at degree of freedom at left margin then the probability for the value of χ^2 will be greater then content in the table is α as shown on top margin $[P(\chi^2 > \chi^2_1) = \alpha]$.

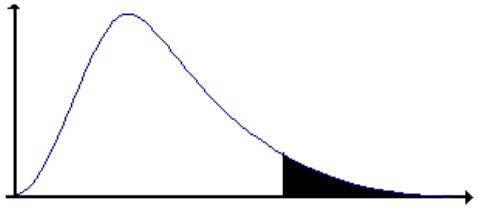
For df = 12 then P (χ^2 > 21.0261) = 0.05; and
for df = 15 the P (χ^2 > 8.5468) = 0.90



α / df	0.975	0.95	0.90	0.10	0.05	0.025	0.01
1	0.00098	0.00393	0.01579	2.70554	3.84146	5.02390	6.63489
2	0.05064	0.10259	0.21072	4.60518	5.99148	7.37778	9.21035
3	0.21579	0.35185	0.58438	6.25139	7.81472	9.34840	11.34488
4	0.48442	0.71072	1.06362	7.77943	9.48773	11.14326	13.27670
5	0.83121	1.14548	1.61031	9.23635	11.07048	12.83249	15.08632
6	1.23734	1.63538	2.20413	10.64464	12.59158	14.44935	16.81187
7	1.68986	2.16735	2.83311	12.01703	14.06713	16.01277	18.47532
8	2.17972	2.73263	3.48954	13.36156	15.50731	17.53454	20.09016
9	2.70039	3.32512	4.16816	14.68366	16.91896	19.02278	21.66605
10	3.24696	3.94030	4.86518	15.98717	18.30703	20.48320	23.20929
11	3.81574	4.57481	5.57779	17.27501	19.67515	21.92002	24.72502
12	4.40378	5.22603	6.30380	18.54934	21.02606	23.33666	26.21696
13	5.00874	5.89186	7.04150	19.81193	22.36203	24.73558	27.68818
14	5.62872	6.57063	7.78954	21.06414	23.68478	26.11893	29.14116
15	6.26212	7.26093	8.54675	22.30712	24.99580	27.48836	30.57795
16	6.90766	7.96164	9.31224	23.54182	26.29622	28.84532	31.99986
17	7.56418	8.67175	10.08518	24.76903	27.58710	30.19098	33.40872
18	8.23074	9.39045	10.86494	25.98942	28.86932	31.52641	34.80524
19	8.90651	10.11701	11.65091	27.20356	30.14351	32.85234	36.19077
20	9.59077	10.85080	12.44260	28.41197	31.41042	34.16958	37.56627
21	10.28291	11.59132	13.23960	29.61509	32.67056	35.47886	38.93223
22	10.98233	12.33801	14.04149	30.81329	33.92446	36.78068	40.28945
23	11.68853	13.09051	14.84795	32.00689	35.17246	38.07561	41.63833
24	12.40115	13.84842	15.65868	33.19624	36.41503	39.36406	42.97978
25	13.11971	14.61140	16.47341	34.38158	37.65249	40.64650	44.31401
26	13.84388	15.37916	17.29188	35.56316	38.88513	41.92314	45.64164
27	14.57337	16.15139	18.11389	36.74123	40.11327	43.19452	46.96284
28	15.30785	16.92788	18.93924	37.91591	41.33715	44.46079	48.27817
29	16.04705	17.70838	19.76774	39.08748	42.55695	45.72228	49.58783
30	16.79076	18.49267	20.59924	40.25602	43.77295	46.97922	50.89218
40	24.43306	26.50930	29.05052	51.80504	55.75849	59.34168	63.69077
50	32.35738	34.76424	37.68864	63.16711	67.50481	71.42019	76.15380
60	40.48171	43.18797	46.45888	74.39700	79.08195	83.29771	88.37943
80	57.15315	60.39146	64.27784	96.57820	101.87947	106.62854	112.32879
100	74.22188	77.92944	82.35813	118.49800	124.34210	129.56125	135.80689
120	91.57260	95.70462	100.62363	140.23256	146.56731	152.21133	158.95003

F Distribution with $\alpha = 0.05$

The table shows that at df of numerator as shown at top margin and df of denominator as shown at left margin, then the value of F will be greater than the content in the table is α [$P(F > F_1) = \alpha$].



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20


Kanopi FEBUI
Unity in Development