

KIỂM TRA

Thời gian: 60 phút.

Sinh viên được xem tài liệu offline, không sử dụng internet để làm bài

Lưu ý:

+ Sinh viên phải bật camera trong suốt thời gian làm bài và gửi ultraview cho giáo viên

+ Nghiêm cấm các trường hợp sử dụng điện thoại trao đổi, gửi bài nếu bị phát hiện sẽ không thu bài

Bài 1 (3đ): Cho dãy số: 18, 22, 24, 26, 32, 35, 41, 43, 48, 52, 65

- Làm mịn dữ liệu bằng phương pháp chia giỏ theo độ sâu với $N = 3$ giá trị trung vị)
Chọn 1 trong 2 câu b hoặc c
- Chuẩn hóa dãy trên về đoạn $[0, 1]$ theo phương pháp min - max.
- Hãy sử dụng phương pháp z-score để chuẩn hóa giá trị 43.

Bài 2 (4đ): Cho CSDL sau và $\text{minsupp} = 50\%$

TID	Item
1	A, B, C
2	A, C
3	A, D
4	B, C
5	A, B, C, D

- Tìm tất cả các tập phổ biến bằng phương pháp Apriori
- Tìm tất cả tập phổ biến đóng
- Tìm tất cả tập phổ biến tối đại
- Xây dựng cây FP-Tree cho CSDL giao dịch trên
- Tính độ phổ biến và độ tin cậy của luật $A \rightarrow C$

Bài 3 (3đ): Liệu sở thích xem phim (hành động hoặc tình cảm) có phụ thuộc vào giới tính (nam hoặc nữ)?

Khảo sát một nhóm người về sở thích xem phim của họ, dữ liệu được thu thập như sau:

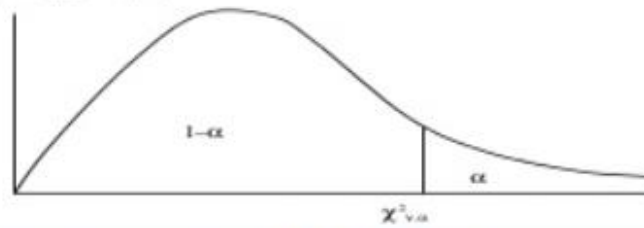
	Hành động	Tình cảm
Nam	40	10
Nữ	15	35

BẢNG TRA PHẦN PHỐI CHI-SQUARE

(Cho $\alpha=0.1$, bậc tự do =6, giá trị Chi-square α sẽ là 10.64.

Ý nghĩa: $P(\text{Chi-square} > \text{Chi-square}_{\alpha}) = \alpha$

Mật độ XS



Bậc tự do v	Chi-Square Alpha									
	0.995	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.005
1	3.93E-05	1.57E-04	9.82E-04	3.93E-03	0.01579	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.610	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81	18.55
7	0.989	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80