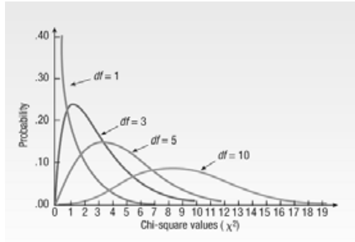
Nonparametric Methods: Chi-square Distribution

Chapter 17

Characteristics of the Chi-Square Distribution

The major characteristics of the chi-square distribution are:

- It is positively skewed.
- It is non-negative.
- It is based on degrees of freedom.
- When the degrees of freedom change, a new distribution is created.



Nonparametric Test

- For the tests studied before, we assume the populations follow the normal probability distribution.
- There are tests available in which no assumption regarding the shape of the population is necessary. These tests are referred to as nonparametric tests.
- The assumption of a normal population is not necessary.

Goodness-of-Fit Test
適 合 度 檢 定

Equal Expected Frequencies

Goodness-of-Fit Test: Equal Expected Frequencies

- The purpose of the goodness-of-fit test is to compare an observed distribution to an expected distribution.
  - Equal Expected Frequencies

    某冰品公司生產三種不同口味的冰棒，想研究顧客偏好三種冰棒的比例是否相同？

  - Unequal Expected Frequencies

    某冰品公司生產三種不同口味的冰棒，想研究顧客偏好三種冰棒的比例是否為 20%, 30%, 50%？

Goodness-of-Fit Test: Equal Expected Frequencies

- H_0: There is no difference between the observed and expected frequencies.
- H_1: There is a difference between the observed and the expected frequencies.

- Notations:
  - f_0 : observed frequencies
  - f_e : expected frequencies

Goodness-of-fit Test: Equal Expected Frequencies

The test statistic is:

$$\chi^2 = \sum\left[\frac{(f_o - f_e)^2}{f_e}\right]$$

The critical value is a chi-square value with (k-1) degrees of freedom, where k is the number of categories

Goodness-of-Fit Example

Ms. Jan Kilpatrick is the marketing manager for a manufacturer of sports cards. She plans to begin selling a series of cards with pictures and playing statistics of former Major League Baseball players. One of the problems is the selection of the former players. At a baseball card show at Southwyck Mall last weekend, she set up a booth and offered cards of the following six Hall of Fame baseball players: Tom Seaver, Nolan Ryan, Ty Cobb, George Brett, Hank Aaron, and Johnny Bench.

Player	Cards Sold
Tom Seaver	13
Nolan Ryan	33
Ty Cobb	14
George Brett	7
Hank Aaron	36
Johnny Bench	17
Total	120

Goodness-of-Fit Example

At the end of the day she sold a total of 120 cards. The number of cards sold for each old-time player is shown in the table on the right. Can she conclude the sales are not the same for each player? Use 0.05 significance level.

Player	Cards Sold
Tom Seaver	13
Nolan Ryan	33
Ty Cobb	14
George Brett	7
Hank Aaron	36
Johnny Bench	17
Total	120

Goodness-of-Fit Example

Step 1: State the null hypothesis and the alternate hypothesis.

H_0: there is no difference between f_o and f_e
H_1: there is a difference between f_o and f_e

Step 2: Select the level of significance.

α = 0.05 as stated in the problem

Goodness-of-Fit Example

Step 3: Select the test statistic.

The test statistic follows the chi-square distribution, designated as χ^2

CHI-SQUARE TEST STATISTIC

$$\chi^2 = \Sigma\left[\frac{(f_o - f_e)^2}{f_e}\right]$$

Goodness-of-Fit Example

Step 4: Formulate the decision rule.

Reject H_0 if $\chi^2 > \chi^2_{\alpha,k-1}$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{\alpha,k-1}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{.05,6-1}$$

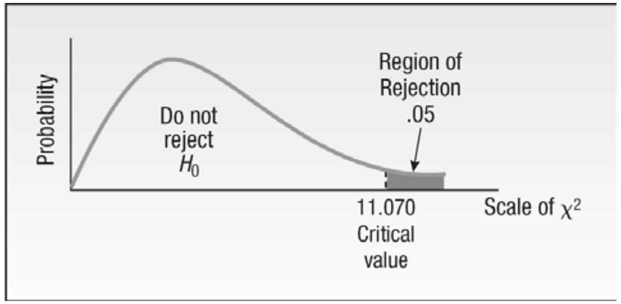
$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{.05,5}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > 11.070$$

A Portion of the Chi-Square Table

Degrees of Freedom	Right-Tail Area			
df	.10	.05	.02	.01
1	2.706	3.841	5.412	6.635
2	4.605	5.991	7.824	9.210
3	6.251	7.815	9.837	11.345
4	7.779	9.488	11.668	13.277
5	9.236	11.070	13.388	15.086

Goodness-of-Fit Example



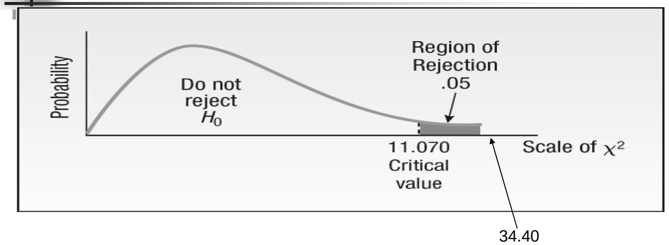
Goodness-of-Fit Example

Step 5: Compute the value of the Chi-square statistic and make a decision

$$\chi^2 = \sum\left[\frac{(f_o - f_e)^2}{f_e}\right]$$

Baseball Player	f_o	f_e	(1) $(f_o - f_e)$	(2) $(f_o - f_e)^2$	(3) $\frac{(f_o - f_e)^2}{f_e}$
Tom Seaver	13	20	−7	49	49/20 = 2.45
Nolan Ryan	33	20	13	169	169/20 = 8.45
Ty Cobb	14	20	−6	36	36/20 = 1.80
George Brett	7	20	−13	169	169/20 = 8.45
Hank Aaron	36	20	16	256	256/20 = 12.80
Johnny Bench	17	20	−3	9	9/20 = 0.45
			0		34.40
		Must be		χ^2	

Goodness-of-Fit Example



The computed χ^2 of 34.40 is in the rejection region, beyond the critical value of 11.070. The decision, therefore, is to reject H_0 at the .05 level . Conclusion: The difference between the observed and the expected frequencies is not due to chance. Rather, the differences between f_0 and f_e and are large enough to be considered significant. It is unlikely that card sales are the same among the six players.

Goodness-of-Fit Example

Conclusion:

The difference between the observed and the expected frequencies is not due to chance. Rather, the differences between f_0 and f_e and are large enough to be considered significant. It is unlikely that card sales are the same among the six players.

Limitations of Chi-Square

- If there is an unusually small expected frequency in a cell, chi-square might result in a wrong conclusion. Two generally accepted policies regarding small cell frequencies are:
  - If there are only two cells, the expected frequency in each cell should be at least 5.
  - For more than two cells, chi-square should not be used if more than 20 % of the fe cells have expected frequencies less than 5.

Examples:

Individual	f0	fe
Literate	643	642
illiterate	7	6

OK !!

Level of Management	f0	fe
Foreman	30	32
Supervisor	110	113
Manager	86	87
Middle Management	23	24
Assistant Vice President	5	2
Vice President	5	4
Senior Vice President	4	1

Problem!!

Resolution

- We can combine the 3 vice president categories.

Level of Management	f0	fe
Foreman	30	32
Supervisor	110	113
Manager	86	87
Middle Management	23	24
President	14	7

OK!!

Goodness-of-Fit Test

Unequal Expected Frequencies

Goodness-of-Fit Test: Unqual Expected Frequencies

- H_0: There is no difference between the observed and expected frequencies.
- H_1: There is a difference between the observed and the expected frequencies.
- Notations:
  - f_0 : observed frequencies
  - f_e : expected frequencies

Goodness-of-Fit Test: Unequal Expected Frequencies - Example

The American Hospital Administrators Association (AHAA) reports the following information concerning the number of times senior citizens are admitted to a hospital during a one-year period.

40 percent are not admitted; 30 percent are admitted once; 20 percent are admitted twice, and the remaining 10 percent are admitted three or more times.

Goodness-of-Fit Test: Unequal Expected Frequencies - Example

A survey of 150 residents of Bartow Estates, a community devoted to active seniors located in central Florida, revealed 55 residents were not admitted during the last year, 50 were admitted to a hospital once, 32 were admitted twice, and the rest of those in the survey were admitted three or more times.

Can we conclude the survey at Bartow Estates is consistent with the information suggested by the AHAA? Use the .05 significance level.

Goodness-of-Fit Test: Unequal Expected Frequencies - Example

Step 1: State the null hypothesis and the alternate hypothesis.

H_0: There is no difference between local and national experience for hospital admissions.

H_1: There is a difference between local and national experience for hospital admissions.

Step 2: Select the level of significance.

α = 0.05 as stated in the problem

Step 3: Select the test statistic.

The test statistic follows the chi-square distribution, designated as χ^2

CHI-SQUARE TEST STATISTIC $$\chi^2 = \Sigma\left[\frac{(f_o - f_e)^2}{f_e}\right]$$

Goodness-of-Fit Test: Unequal Expected Frequencies - Example

Step 4: Formulate the decision rule.

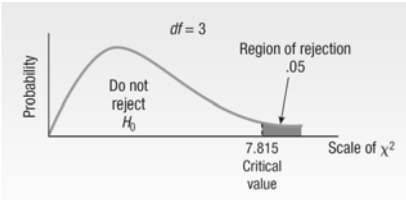
Reject H_0 if $\chi^2 > \chi^2_{\alpha,k-1}$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{\alpha,k-1}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{.05,4-1}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{.05,3}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > 7.815$$



Goodness-of-Fit Test: Unequal Expected Frequencies - Example

Distribution stated in the problem

Frequencies observed in a sample of 150 Bartow residents

Expected frequencies of sample if the distribution stated in the Null Hypothesis is correct

Number of Times Admitted	AHAA Percent of Total	Number of Bartow Residents (f_o)	Expected Number of Residents (f_e)
0	40	55	60
1	30	50	45
2	20	32	30
3 or more	10	13	15
Total	100	150	150

Computation of f_e
0.40 X 150 = 60
0.30 X 150 = 45
0.30 X 150 = 30
0.10 X 150= 15

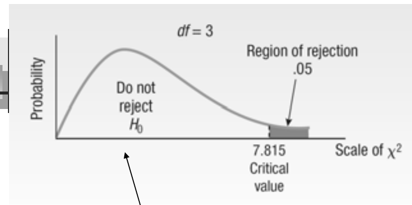
Goodness-of-Fit Test: Unequal Expected Frequencies - Example

Step 5: Compute the value of the Chi-square statistic and make a decision

$$\chi^2 = \sum\left[\frac{(f_o - f_e)^2}{f_e}\right]$$

Number of Times Admitted	(f_o)	(f_e)	$f_o - f_e$	$(f_o - f_e)^2/f_e$
0	55	60	−5	0.4167
1	50	45	5	0.5556
2	32	30	2	0.1333
3 or more	13	15	−2	0.2667
Total	150	150	0	1.3723

Computed χ^2



1.3723

The computed χ^2 of 1.3723 is in the “Do not rejection H_0” region. The difference between the observed and the expected frequencies is due to chance.

We conclude that there is no evidence a difference between the local and national experience for hospital admissions.

Ex. 17.1

- 已知人民對經濟發展的看法認為好的比率為 65%，不好的比率為 20%，沒意見的比率為 15%。現調查 1000 個公民，結果如下表。問人民的看法是否改變？

對經濟發展看法	人數
好	300
不好	600
沒意見	100

Ex. 17.1

1.設立兩個假設：H_0: 看法一樣; H_1: 看法不一樣

2.選擇檢定統計量

因爲有三個類別，是一個多項實驗，因此以卡方分配來做檢定。

3.決定拒絕域或接受域

$\alpha = 0.025$，卡方檢定是右尾檢定。k = 3，自由度 df = k - 1 = 3 - 1 = 2。

查卡方機率值表得臨界值 $\chi^2_{2,0.025} = 7.378$。

Ex. 17.1

- 4. 計算檢定統計量

卡方檢定統計量的計算如下表

看法	O	機率	E=np	(O-E)	$(O-E)^2$	$(O-E)^2/E$
好	300	0.65	1000*0.65=650	-350	122500	188.46
不好	600	0.2	1000*0.2=200	400	160000	800
沒意見	100	0.15	1000*0.15=150	-50	2500	16.67
						1005.13

Ex. 17.1 (1)

由上表知卡方檢定統計量爲: $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 1005.13$

5.下結論

檢定統計量$\chi^2 = 1005.13 > 7.378$，落在拒絕域，因此

拒絕虛無假設。故可下結論說人民對經濟發展的看法已經改變。

Ex. 17.2 (1)

- 假設統計學的期中分數如下

分數	0-20	20-40	40-60	60-80	80-100	
人數	8	14	47	20	11	n=100

請問該分配是否為常態分配？

Ex. 17.2 (2)

1.設立兩個假設：H_0: 爲常態分配; H_1: 不爲常態分配

2.選擇檢定統計量

因爲有5個類別，是一個多項實驗，因此以卡方分配來做檢定。

3.決定拒絕域或接受域

α = 0.025，k = 5，自由度 df = k-1= 5-1-2 = 2

(其中 m = 2 是因爲利用樣本資料估計兩個未知參數 μ 及 σ^2)。

查卡方機率值表得臨界值 $\chi^2_{2,0.025}$ = 7.378。

Ex. 17.2 (3)

4.計算檢定統計量

因常態分配的參數 μ 未知故以 $\bar{X}$ 估計之

$$\bar{X} = \frac{10*8+30*14+50*47+70*20+90*11}{100} = 52.4$$

另一參數 σ 未知故以 S 估計之

$$S^2 = \frac{10^2*8+30^2*14+50^2*47+70^2*20+90^2*11-100*52.4^2}{100-1} = 438.48$$

故可得 S = 20.94

Ex. 17.2 (4)

故可得 $P(X < 20) = P(Z < \frac{20-52.4}{20.94}) = P(Z < -1.547) = 0.061$

$P(20 < X < 40) = P(\frac{20-52.4}{20.94} < Z < \frac{40-52.4}{20.94})$

$= P(-1.547 < Z < -0.592) = 0.217$

另外三組的機率值如下

Ex. 17.2 (5)

分數	20以下	20-40	40-60	60-80	80以上
O	8	14	47	20	11
Z=(U-52.4)/20.94	-1.547	-0.592	0.363	1.318	∞
Z=(L-52.4)/20.94	-∞	-1.547	-0.592	0.363	1.318
機率值P	0.061	0.217	0.363	0.266	0.093
E(np)	6.1	21.7	36.3	26.6	9.3
$(O-E)^2$	3.61	59.29	114.49	43.56	2.89

Ex. 17.2 (6)

由上知 $\bar{X}$ = 52.4，S = 20.94，故可求得各組的上下限的 Z 值。

結果如表第三列第四列。由各組的上下限Z值可求各組的機率如表第五列。第六列爲期望次數即 E(np)。第七列爲觀察次數與期望次數的平方。因此可得

$$\chi^2 = \frac{3.61}{6.1} + \frac{59.29}{21.7} + \frac{114.49}{36.3} + \frac{43.56}{26.6} + \frac{2.89}{9.3} = 8.43$$

Ex. 17.2 (7)

5.下結論

檢定統計量 χ^2 = 8.43 > 臨界值 $\chi^2_{5-1-2,0.025}$ = 7.378 ，落在拒絕域，因此拒絕虛無假設，故可下結論說統計學期中考成績不爲常態分配。

資料是否爲常態分配可繪製直方圖來觀察，但直方圖不如統計數字來得精確。

列聯表分析Contingency Table Analysis 獨立性檢定Test of Independence

Contingency Tables

A CONTINGENCY TABLE is a table used to classify sample observations according to two or more identifiable characteristics

Contingency Tables

E.g. A survey of 150 adults classified each as to gender and the number of movies attended last month. Each respondent is classified according to two criteria—the number of movies attended and gender.

Is gender related to the movies attended ?

	Gender		
Movies Attended	**Men**	**Women**	**Total**
0	20	40	60
1	40	30	70
2 or more	10	10	20
Total	70	80	150

Contingency Tables - Example

A sample of executives were surveyed about their loyalty to their company. One of the questions was, "If you were given an offer by another company equal to or slightly better than your present position, would you remain with the company or take the other position?" The responses of the 200 executives in the survey were cross-classified with their length of service with the company.

Is loyalty related to the length of employment ?

	Length of Service				
Loyalty	**Less than 1 Year, B_1**	**1–5 Years, B_2**	**6–10 Years, B_3**	**More than 10 Years, B_4**	**Total**
Would remain, A_1	10	30	5	75	120
Would not remain, A_2	25	15	10	30	80
	35	45	15	105	200

Contingency Table Analysis

A contingency table is used to investigate whether two traits or characteristics are related. Each observation is classified according to two criteria. We use the usual hypothesis testing procedure.

- The *degrees of freedom* is equal to:

  (number of rows-1)(number of columns-1).

- The *expected frequency* is computed as:

EXPECTED FREQUENCY

$$f_e = \frac{(\text{Row total})(\text{Column total})}{\text{Grand total}}$$

Contingency Analysis

We can use the chi-square statistic to formally test for a relationship between two nominal-scaled variables. To put it another way, Is one variable *independent* of the other?

- Ford Motor Company operates an assembly plant in Dearborn, Michigan. The plant operates three shifts per day, 5 days a week. The quality control manager wishes to compare the quality level on the three shifts. Vehicles are classified by quality level (acceptable, unacceptable) and shift (day, afternoon, night). Is there a difference in the quality level on the three shifts? That is, is the quality of the product related to the shift when it was manufactured? Or is the quality of the product independent of the shift on which it was manufactured?

Contingency Analysis

- A sample of 100 drivers who were stopped for speeding violations was classified by gender and whether or not they were wearing a seat belt. For this sample, is wearing a seatbelt related to gender?
- Does a male released from federal prison make a different adjustment to civilian life if he returns to his hometown or if he goes elsewhere to live? The two variables are adjustment to civilian life and place of residence. Note that both variables are measured on the nominal scale.

Contingency Analysis - Example

The Federal Correction Agency is investigating the last question cited above: Does a male released from federal prison make a different adjustment to civilian life if he returns to his hometown or if he goes elsewhere to live? To put it another way, is there a relationship between adjustment to civilian life and place of residence after release from prison? Use the .01 significance level.

Contingency Analysis - Example

The agency's psychologists interviewed 200 randomly selected former prisoners. Using a series of questions, the psychologists classified the adjustment of each individual to civilian life as outstanding, good, fair, or unsatisfactory. The classifications for the 200 former prisoners were tallied as follows. Joseph Camden, for example, returned to his hometown and has shown outstanding adjustment to civilian life. His case is one of the 27 tallies in the upper left box (circled).

Residence after Release from Prison	Adjustment to Civilian Life			
	Outstanding	Good	Fair	Unsatisfactory
Hometown	卌 卌 卌 卌 卌 ||	卌 卌 卌 卌 卌 卌 卌	卌 卌 卌 卌 卌 卌 |||	卌 卌 卌 卌 卌
Not hometown	卌 卌 |||	卌 卌 卌	卌 卌 卌 卌 卌 ||	卌 卌 卌 卌 卌

Contingency Analysis - Example

Residence after Release from Prison	Adjustment to Civilian Life				
	Outstanding	Good	Fair	Unsatisfactory	Total
Hometown	27	35	33	25	120
Not hometown	13	15	27	25	80
Total	40	50	60	50	200

Contingency Analysis - Example

Step 1: State the null hypothesis and the alternate hypothesis.

H_0: There is no relationship between adjustment to civilian life and where the individual lives after being released from prison.

H_1: There is a relationship between adjustment to civilian life and where the individual lives after being released from prison.

Step 2: Select the level of significance.

α = 0.01 as stated in the problem

Contingency Analysis - Example

Step 3: Select the test statistic.

The test statistic follows the chi-square distribution, designated as χ^2

CHI-SQUARE TEST STATISTIC $\chi^2 = \sum\left[\frac{(f_o - f_e)^2}{f_e}\right]$

Contingency Analysis - Example

Step 4: Formulate the decision rule.

Reject H_0 if $\chi^2 > \chi^2_{\alpha,(r-1)(c-1)}$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{\alpha,(2-1)(4-1)}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{.01,(1)(3)}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > \chi^2_{.01,3}$$

$$\sum\left[\frac{(f_o - f_e)^2}{f_e}\right] > 11.345$$

Computing Expected Frequencies (f_e)

EXPECTED FREQUENCY $f_e = \frac{\text{(Row total)(Column total)}}{\text{Grand total}}$

$\frac{(120)(50)}{200}$

Residence after Release from Prison	Outstanding f_o	Outstanding f_e	Good f_o	Good f_e	Fair f_o	Fair f_e	Unsatisfactory f_o	Unsatisfactory f_e	Total f_o	Total f_e
Hometown	27	24	35	30	33	36	25	30	120	120
Not hometown	13	16	15	20	27	24	25	20	80	80
Total	40	40	50	50	60	60	50	50	200	200

Adjustment to Civilian Life

Must be equal

$\frac{(80)(50)}{200}$

Must be equal

Computing the Chi-square Statistic

Residence after Release from Prison	Outstanding f_o	Outstanding f_e	Good f_o	Good f_e	Fair f_o	Fair f_e	Unsatisfactory f_o	Unsatisfactory f_e	Total f_o	Total f_e
Hometown	27	24	35	30	33	36	25	30	120	120
Not hometown	13	16	15	20	27	24	25	20	80	80
Total	40	40	50	50	60	60	50	50	200	200

Adjustment to Civilian Life

$$\chi^2 = \Sigma\left[\frac{(f_o - f_e)^2}{f_e}\right]$$

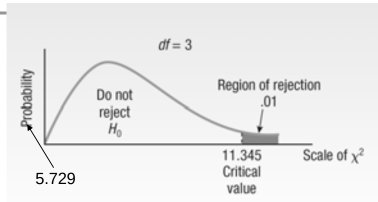
Starting with the upper left cell:

$$\chi^2 = \frac{(27-24)^2}{24} + \frac{(35-30)^2}{30} + \frac{(33-36)^2}{36} + \frac{(25-30)^2}{30} + \frac{(13-16)^2}{16} + \frac{(15-20)^2}{20} + \frac{(27-24)^2}{24} + \frac{(25-20)^2}{20}$$

$$= 0.375 + 0.833 + 0.250 + 0.833 + 0.563 + 1.250 + 0.375 + 1.250$$

$$= 5.729$$

Conclusion



The computed χ^2 of 5.729 is in the "Do not rejection H_0" region. The null hypothesis is not rejected at the .01 significance level.

We conclude there is no evidence of a relationship between adjustment to civilian life and where the prisoner resides after being released from prison. For the Federal Correction Agency's advisement program, adjustment to civilian life is not related to where the ex-prisoner lives.

Ex. 17.3 (1)

- 根據下表，試問客戶年齡與喜好車型是否有關？

車型	30歲以下	31歲以上	合計
Premacy	108	74	182
TRIBUTE2.0/3.0	45	36	81
MPV	90	126	216
合計	243	236	479

Ex. 17.3 (2)

1.設立兩個假設：H_0: 客戶年齡與喜好車型無關; H_1: 客戶年齡與喜好車型有關

2.選擇檢定統計量

以卡方分配來做檢定。

3.決定拒絕域或接受域

α = 0.05，上表是一個 3 * 2 的列聯表，因此自由度 df = (c-1)(r-1) = (3-1)(2-1) = 2

查卡方機率值表得臨界值 $\chi^2_{2,0.05} = 5.991$。

Ex. 17.3 (3)

4.計算檢定統計量

依前面 $\hat{E}_{ij} = \dfrac{\text{i 列總和} * \text{j 行總和}}{\text{樣本數}}$ 公式可以計算理論次數如下表

括弧內的理論次數

車型/年齡	30歲以下	31歲以上	合計
Premacy	108(92.33)	74(89.67)	182
TRIBUTE2.0/3.0	45(41.09)	36(39.91)	81
MPV	90(109.58)	126(106.42)	216
合計	243	236	479

Ex. 17.3 (4)

接著計算檢定統計量 $\chi^2 = \dfrac{(108-92.33)^2}{92.33} + \dfrac{(74-89.67)^2}{89.67} + ... + \dfrac{(126-106.42)^2}{106.42} = 13.254$

5.下結論

因檢定統計量 χ^2 =13.254 大於臨界值 $\chi^2_{2,0.05} = 5.991$，落在拒絕域

故拒絕虛無假設，結論爲客戶年齡與車型喜好有關。事實上，三十歲以下的客戶經濟能力較差，家庭成員較少，故較低價車體小的Premacy 爲年輕族較爲喜愛的休旅車。

Ex. 17.4 (1)

- 調查 398 個工人的婚姻狀況如下表，試檢定婚姻狀況與離職意願是否有關 (α = 0.01)?

離職意願	婚姻狀況	婚姻狀況	合計
	未婚	已婚	
低	80	61	141
中	94	24	118
高	108	31	139
合計	282	116	398

Ex. 17.4 (2)

設立兩個假設：H_0: 婚姻狀況與離職意願無關

H_1: 婚姻狀況與離職意願有關

計算理論次數可得

$\hat{E}_{11}$ =398*141/398*282/398=99.9 ，$\hat{E}_{12}$ = 398*141/398*116/398=41.1

$\hat{E}_{21}$ =398*118/398*282/398=83.6，$\hat{E}_{22}$ =398*118/398*116/398=34.4

$\hat{E}_{31}$ =398*139/398*282/398=98.5，$\hat{E}_{32}$ =398*139/398*116/398=40.5

Ex. 17.4 (3)

計算檢定統計量

$$\chi^2=\frac{(80-99.9)^2}{99.9}+\frac{(61-41.1)^2}{41.1}+\frac{(94-83.6)^2}{83.6}+\frac{(24-34.4)^2}{34.4}$$

$$+\frac{(108-98.5)^2}{98.5}+\frac{(31-40.5)^2}{40.5}=21.18$$

因 $\chi^2=21.18$ 大於臨界值 $\chi^2_{(3-1)(2-1),0.01}=9.21$

故拒絕虛無假設。結論爲婚姻狀況與離職高低有關。

Exercises

- 1,3,5,7,9,13,15,17,19,21,25,27