Handout Mainstat(2) ? yes. I

ex.'s. for χ^2 test for Contingency tables

ex.1.

年收入 Annual income (US dollars)

length of time (years) with the same company (年資) (Eij)	< 8000	8000–15000	> 15000	
< 5	50 (37.5)	75 (75)	25 (37.5)	150
5 – 10	25 (25)	50 (50)	25 (25)	100
10 or more	25 (37.5)	75 (75)	50 (37.5)	150
	100	200	100	400

Test H_0: 年資 ⊥ 年收入 at $\alpha = 0.05$

$$\Rightarrow \chi^2 = \frac{(50-37.5)^2}{37.5} + \frac{(75-75)^2}{75} + \cdots + \frac{(50-37.5)^2}{37.5}$$

$$= 16.66 > \chi^2_{4,0.05} = 9.488$$

$\Rightarrow$ Reject H_0 at level 0.05.

I.e. there is sig. evidence that length of service with a company is not indep. of annual income.
(i.e. annual income is related to the length of service at the company)

ex.2.

Pension plan

Job type	A	B	C	
full time	160 (136)	140 (136)	40 (68)	340
part time	40 (64)	60 (64)	60 (32)	160
	200	200	100	500

Test H_0: preference for pension plan is indep. of job type.
At $\alpha = 0.05$, H_0 is rejected, i.e. one's job type does influence his preference on pension plan.
since $\chi^2 = 49.63 > \chi^2_{2,0.05} = 5.99$

ex. 3. Y = # of days spent exercising per week

X = Gender

	0-1	2-3	4-5	6-7	
M	40	53	26	6	125
F	34	68	37	11	150
	74	121	63	17	275

Test $H_0: X \perp Y$ at $\alpha = 0.05$

$\Rightarrow \chi^2 = 3.49 < \chi^2_{3, 0.05} = 7.815$

$\Rightarrow$ H_0 is not rejected at level 0.05

I.e. There is not enough evidence to conclude that the # of days spent exercising per week is related to one's gender.

ex 4. the effectiveness of two commercials A, B

see A or B $\xrightarrow{2 \text{ hrs}}$ interview. ?, remember the k-p

Response of the audience

commercial	Don't Remember	Remember seeing	Remember key point	
A	19 (22.93)	24 (27.73)	37 (29.33)	80
B	24 (20.07)	28 (24.27)	18 (25.6)	70
	43	52	55	150

$\Rightarrow \chi^2 = 6.816 > \chi^2_{2, 0.05} = 5.99$

$\Rightarrow$ H_0 is rejected at level 0.05

I.e. There is sig. evidence to conclude that the audience response differently to these two commercials.

Ex.(1) Calulus Midterm w/ $n = 63$, $\text{mean} = 31.6$; $\text{s.e.} = 21.6$
$\hat{\mu}$ " $\hat{\sigma}$ "

H_0: grade $\sim N(\mu, \sigma^2)$ v.s. Not H_0.

Data score	O_i	x	$z = \frac{x - 31.6}{21.6}$	$\Phi(z)$	$\hat{p}_i \times n = E_i$	
<10	11	"			0.1587	9.99
10-19	10	10	-1	0.1587	0.1359	8.56
20-29	10	20	-0.56	0.2946	0.1775	11.16
30-39	10	30	-0.07	0.4721	0.1796	11.31
40-59	14	40	0.39	0.6517	0.2532	15.95
≥60	8	60	1.31	0.9049	0.0951	5.99
				1		

$$\Rightarrow \chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i} = \frac{(11-9.99)^2}{9.99} + \frac{(10-8.56)^2}{8.56} + \cdots + \frac{(8-5.99)^2}{5.99} = 1.52$$

$$< \chi^2_{6-1-2,0.05} = \chi^2_{3,0.05} = 7.815$$
$$(\chi^2_{3,0.1} = 6.251)$$

$\Rightarrow$ H_0 is Not rejected at level 0.05 (or 0.1)

I.e. Normal assumption is not violated, hence, suitable for analyzing this data set.

Ex.(2) A: [90, 100], B: [75, 90), C: [60, 75), D: [50, 60), E: [0, 50)

$n = 52$

A	B	C	D	E
3	12	17	14	6

H_0: grade $\sim N(65, 10^2)$?

O_i	x	$z = \frac{x-65}{10}$	$\Phi(z)$	$\hat{p}_i$	$\times n = E_i$		$(O_i - E_i)^2/E_i$
6				0.0559	2.91 (<5)	14.44	2.14
14	50	-1.5	0.0559	0.2217	11.53		
17	60	-0.5	0.2776	0.5637	29.3		5.17
12	75	1	0.8413	0.1525	7.93	8.25	5.52
3	90	2.5	0.9938	0.0062	0.3224 (<5)		
	100	3.5	1				

(20 = 6 + 14; 15 = 12 + 3)

$\chi^2 = 12.83$

$\chi^2_{3-1,0.05} = \chi^2_{2,0.05} = 5.991$

$\Rightarrow$ Reject H_0 at level 0.05

(Note: If no combining.
$\chi^2 = 32.6 > \chi^2_{5-1,0.05} = \chi^2_{4,0.05} = 9.488$)

I.e. Normal assumption seems to be violated with this data.

Ex(3). Y = # of typos in a page of a certain book.

Ho: $Y \sim P(\lambda)$

Data:

i: # of typos	(O_i) frequency	$\hat{P}_i = \hat{P}(Y=i) = \frac{e^{-\hat{\lambda}}\hat{\lambda}^i}{i!}$	$\times n = E_i$
0	18	0.0498	21.9
1	53	0.1494	65.7
2	103	0.2240	98.6
3	107	0.2240	98.6
4	82	0.1680	73.9
5	46	0.1008	44.4
6	18	0.0504	22.2
7	10	0.0216	9.5
8	3 < 2	0.0081	3.6 > 5.3
9	1	0.0038	1.7
	440 = n.		

k=9

Note $\hat{\lambda}_{MLE} = \bar{Y} = (18\times0 + 53\times1 + 103\times2 + 107\times3 + \cdots + 2\times8 + 1\times9)/440 \approx 3.$

$$\Rightarrow \chi^2 = \frac{(18-21.9)^2}{21.9} + \frac{(53-65.7)^2}{65.7} + \cdots + \frac{(3-5.3)^2}{5.3} = 6.83$$

$$\chi^2_{9-1-1,\,0.05} = \chi^2_{7,\,0.05} = 14.067$$

$\Rightarrow$ Ho is Not rejected at level 0.05.

I.e. Poisson dist. provides a "good fit" to this data set.

Ex(4).

Ho: $X \sim N(\mu, 10^2)$ $\rightsquigarrow$ $\chi^2_{K-1-1,\alpha}$

Ho: $X \sim N(\mu, \sigma^2)$ $\rightsquigarrow$ $\chi^2_{K-1-2,\alpha}$

Ho: $X \sim$ Gamma$(1, \beta)$ $\rightsquigarrow$ $\chi^2_{K-1-1,\alpha}$.

etc.