18.1 General consideration *P.498*

1 - 3. General procedure (self-study) *P.498-500*

4. Calculate the expected class frequencies

   expected / fitted freq. $=$ total freq. $\times$ fitted prob.

   $f_e = N \times \hat{p}$

5. Compare the class frequencies

   For evaluating model, one common procedure is to compare the observed frequency, f_o, with the corresponding expected frequency, f_e.

   Absolute discrepancy $= | f_o - f_e |$

   (a) large values indicate a poor fit, small values indicate a good fit.

   (b) But at this level, the judgment is quite subjective.

6. At a more advanced level, chi - square (χ^2) test will be used to test the goodness - of - fit of the theoretical distributions.

18.2 Fitting a discrete uniform distribution *P.501*

Exercise 18.2

4. $N = 300$

Score	freq., f_o	$\hat{p}$	$f_e = N \times \hat{p}$	$\| f_o - f_e \|$
1	47	1/6	50	3
2	52			
3	48			
4	57			
5	56			
6	40			
	$\Sigma = 300$			

$\because$ discrepancies are not large

$\therefore$ the die is fair.

9. $N = 160$

Company	Value of sales ($m)	$\hat{p}$	Expected sales ($m)	$\| f_o - f_e \|$
A	45			
B	41			
C	19			
D	30			
E	25			
	$\Sigma = 160$	$\Sigma = 1$	$\Sigma = 160$	

$\because$ discrepancies are small

$\therefore$ the data support the claim.

10. $N = 150$

(a)

Colour	freq., f_o	$\hat{p}$	$f_e = N \times \hat{p}$	$\| f_o - f_e \|$
R	69	10/25	60	9
Y	35			
B	24			
G	22			
	$\Sigma = 150$			

Do it yourself.

(b)

Colour	freq., f_o	$\hat{p}$	$f_e = N \times \hat{p}$	$\| f_o - f_e \|$
R	390	10/25	400	10
Y	263			
B	212			
G	135			
	$\Sigma = 1000$			

Do it yourself.

18.3 Fitting a Poisson distribution

P.506

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{for } x = 0, 1, 2, \cdots$$

Cases 1. λ is known ⇒ straight forward
2. λ is unknown ⇒ find the sample mean of the empirical distribution.

Exercise 18.3

2. (a) Let X be no. of calls per hr.

$$P(X = x) = \frac{2^x e^{-2}}{x!}$$

x	fo	$\hat{p}$	fe	| fo − fe |
0	17	0.1353	16.2	0.8
1	31			
2	33			
3	20			
4	8			
5	6			
6 or more	5			
	Σ = 120	Σ = 1		

(b) Not agree.

5. (a) mean $= \dfrac{0\times18+1\times28+2\times19+3\times9+4\times4+5\times2}{80}$

=

(b) Let X be no. of faults

$$P(X = x) = \frac{1.4875^x e^{-1.4875}}{x!}$$

x	fo	$\hat{p}$	fe = N × $\hat{p}$	| fo − fe |
0	18	0.2259	18.1	0.1
1	28			
2	19			
3	9			
4	4			
5 or more	2			
	Σ = 80	Σ = 1		

∵ discrepancies are small
∴ support

9. (a) sample mean =

sample variance =

(b) Do it yourself

(c) $\text{Po}(0.9769) = \dfrac{0.9769^x e^{-0.9769}}{x!}$

(d)

No. of animals	fo	p̂	fe	| fo − fe |
0	51	0.3765	48.9	2.1
1	47			
2	20			
3	9			
4	2			
5 or more	1			
	Σ = 130	Σ = 1		

∵ discrepancy for 6th cell are quite large

∴ do not fit.

(d) P (2 adjacent regions have no animal) = P(X=0) · P(X=0)

=

=

18.4 Fitting a Binomial distribution [*B (n , p)* or *b (x ; n , p)*] *P.513*

$P(X = x) = C_x^n p^x (1-p)^{n-x}$ $x = 0, 1, 2, \ldots, n$

mean = μ = np

variance = σ^2 = np(1 - p)

Cases

1. ***n, p*** **are known**

⇒ **straight forward.**

2. ***n*** **is known, but** ***p*** **is not**

⇒ **by** $\bar{x} = n\hat{p}$ **, then** $\hat{p} = \dfrac{\bar{x}}{n}$ **can be found.**

3. **Both** ***n, p*** **are unknown**

⇒ **by** $\bar{x} = \hat{n}\hat{p}$, $s^2 = \hat{n}\hat{p}(1-\hat{p})$, **then** $\bar{x}$ **and** s^2 **can be calculated from the empirical distribution.**

Exercise 18.4

3. (a) Let X be the no. of red fish in each sample.

If the claim is true, then X ~ B(5, 0.75) and

$P(X = x) = C_x^5 (0.75)^x (0.25)^{5-x}$

x	f_o	p̂	f_e	$\|f_o - f_e\|$
0	4	0.0010	0.2	3.8
1	12			
2	36			
3	67			
4	61			
5	20			

(b) Do it yourself.

5. (a) mean = $\bar{x} = \dfrac{0\times 7+1\times 24+2\times 36+3\times 22+4\times 8+5\times 3}{7+24+36+22+8+3} = 2.09$

(b) Let X be the no. of germinating seeds in the pot.
Suppose X ~ B(5, *p*).

mean = $5p = 2.09$
$p = 0.418$
$\therefore$ $P(X = x) = C_x^5 (0.418)^x (0.582)^{5-x}$

x	f_o	$\hat{p}$	f_e	$\lvert f_o - f_e \rvert$
0	7	0.0668	6.7	0.3
1	24			
2	36			
3	22			
4	8			
5	3			

(c) The expected frequencies fit the observed frequencies extremely well except for the last cell. But the discrepancy for the last cell can be tolerated as its observed frequency is small. Therefore, we can conclude that the binomial distribution is appropriate for the data.

8. (a) mean = $\bar{x} = \dfrac{0\times 117+1\times 140+2\times 130+3\times 63+4\times 30}{480} = 1.4771$

(b) Po(1.4771)

$P(X = x) = \dfrac{e^{-1.4771}(1.4771^x)}{x!}$

x	f_o	$\hat{p}$	f_P	$\lvert f_o - f_p \rvert$
0	117	0.2283	109.6	7.4
1	140			
2	130			
3	63			
4 or more	30			

(c) B(4, *p*).
mean = $4p = 1.4771$
$p = 0.3693$
$\therefore$ $P(X = x) = C_x^4 (0.3693)^x (0.6307)^{4-x}$

x	f_o	$\hat{p}$	f_B	$\lvert f_o - f_B \rvert$
0	117	0.1582	76.0	41.0
1	140			
2	130			
3	63			
4	30			

(d) Poisson distribution fits the data better.

(e) The binomial model B(4, p) is not appropriate because the one-minute intervals cannot be regarded as binomial experiments each with 4 Bernoulli trials.

9.5 Fitting a Normal distribution

P.523

1. If either mean or variance or both are unknown, use $\bar{x}=\frac{\sum fx}{\sum f}$ & $s^2=\frac{1}{n-1}\sum f(x-\bar{x})^2$ to find out the sample mean and variance.

2. Standardize the normal distribution, N(μ, σ^2) $\to$ N(0, 1), by using $z=\frac{x-\mu}{\sigma}$.

Exercise 18.5

3. (a) n = 200

Let X be the weight of firefighter,

X ~ N (70, 6^2)

Weight	X	Standardized class interval (Z - value)	$\hat{p}$	fo	fe = N× $\hat{p}$	| fo − fe |
55 - 60	less than 60	-∞ — -1.67	0.0475	7	9.5	2.5
60 - 65	60 - 65	—		38		
65 - 70	65 - 70	—		62		
70 - 75	70 - 75	—		45		
75 - 80	75 - 80	—		39		
80 - 85	80 or above	—		9		
			Σ = 1	Σ = 200		

(b) Do it yourself.

4. (a) n = 50

Time (mins)	mid - value	f_o
10 - 14	12	5
15 - 19		18
20 - 24		17
25 - 29		7
30 - 34		3

mean =

standard deviation = $\sqrt{\dfrac{\sum f(x-\bar{x})^2}{n-1}}$

(b) (i) Let X be the time taken,

$X \sim N(20.5, 5.175^2)$

Time (mins)	X	Z - value	$\hat{p}$	f_o	$f_e = N \times \hat{p}$	$\lvert f_o - f_e \rvert$
10 - 14	less than 14.5	$-\infty$ —	0.1230	5	6.2	1.2
15 - 19	—	—		18		
20 - 24	—	—		17		
25 - 29	—	—		7		
30 - 34	29.5 or above	— ∞		3		
			$\Sigma = 1$	$\Sigma = 50$		

(ii) Do it yourself.