

6D Testing the Goodness of Fit with Discrete Data

The steps to test goodness of fit with discrete data:

1

2

3

4

5

6

7

8

9

10

11

Discrete uniform distribution: conditions

$P(X = x_r) =$ Number of degrees of freedom =

1. 100 digits between 0 and 9 are selected from a table with the frequencies as shown below.

Digit	0	1	2	3	4	5	6	7	8	9
Frequency	11	8	8	7	8	9	12	9	13	15

Could the digits be from a random number table? Test at the 0.05 level.

[illegible]

Binomial distribution: conditions

$P(X = r) =$

Frequency, $f_r =$

Estimating p from the observed values: $p =$

If p is not estimated by calculation: $v =$

If p is estimated by calculation: $v =$

2. The data in the table are thought to be modelled by a binomial distribution $B(10, 0.2)$. Use the table for the binomial cumulative distribution function to find expected values, and conduct a test to see if this is a good model. Use a 5% significance level.

x	0	1	2	3	4	5	6	7	8
Frequency	12	28	28	17	7	4	2	2	0
$P(X = x)$									
Expected									

x	0	1	2	3	
O_i					
E_i					
$(O - E)^2/E$					

3. A study of the number of girls in families with five children was done on 100 such families. The results are summarised in the following table.

Number of girls (r)	0	1	2	3	4	5
Frequency (f)	13	18	38	20	10	1

It is suggested that the distribution may be modelled by a binomial distribution with $p = 0.5$.

a) Give reasons why this might be so.

b) Test, at the 5% significance level, whether or not a binomial distribution is a good model.

r	0	1	2	3	4	5
O_i						
E_i						

r				
O_i				
E_i				
(O - E)²/E				

4. Look at the data in the previous question. By estimating a suitable value of the parameter, p , carry out a test, at the 5% significance level, to determine whether a binomial distribution is a suitable model for the data.

r	P(r)	E_i
0		
1		
2		
3		
4		
5		

r	O_i	E_i	O²/E
0			
1			
2			
3			
Total			

Poisson distribution: conditions

$P(X = r) =$ Frequency, $f_r =$

Grouping the tail: $P(X \geq n) =$

Estimating λ from the observed values: $\lambda =$

If λ is not estimated by calculation: $v =$

If λ is estimated by calculation: $v =$

5. The numbers of telephone calls arriving at an exchange in six-minute periods were recorded over a period of 8 hours, with the following results.

Number of calls, r	0	1	2	3	4	5	6	7	8
Frequency, f _r	8	19	26	13	7	5	1	1	0

Can these results be modelled by a Poisson distribution? Test at the 5% significance level.

r	$P(X = r)$	Expected frequency
0		
1		
2		
3		
4		
5		
6		
7 or more		

r	O _i	E _i	$(O - E)^2/E$
0			
1			
2			
3			
4			