

## Degree of freedom in Person's Chi-squared test for a Poisson distribution

Consider the goodness of fit hypothesis test

$H_0: X \sim \text{Poisson distribution (assuming four values are observed in the sample)}$

Categories	Counts	Observations
1	$O_1$	$x_{1,1}, x_{1,2}, \dots, x_{1,O_1} = 1$
2	$O_2$	$x_{2,1}, x_{2,2}, \dots, x_{2,O_2} = 2$
3	$O_3$	$x_{3,1}, x_{3,2}, \dots, x_{3,O_3} = 3$
4	$O_4$	$x_{4,1}, x_{4,2}, \dots, x_{4,O_4} = 4$

For this null hypothesis, we first need to get an estimate of  $\lambda$

$$\begin{aligned}\hat{\lambda}_{MLE} &= \bar{X}_n = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \left( \sum_{i=1}^{O_1} x_{1,i} + \sum_{i=1}^{O_2} x_{2,i} + \sum_{i=1}^{O_3} x_{3,i} + \sum_{i=1}^{O_4} x_{4,i} \right) \\ &= \frac{1}{n} (O_1 + 2O_2 + 3O_3 + 4O_4)\end{aligned}$$

Hence, there are two constraints in this case

$$(1) \frac{1}{n} (O_1 + 2O_2 + 3O_3 + 4O_4) = \hat{\lambda}_{MLE} = \hat{\lambda}$$

$$(2) O_1 + O_2 + O_3 + O_4 = n$$

Two of  $O_1, O_2, O_3, O_4$  need to be fixed to satisfy both of these two constraints.

Now consider a counterexample, assuming  $O_1, O_2, O_3$  are free to vary, and let's try to solve  $O_4$ :

From (1)

$$O_4 = \frac{n\hat{\lambda} - (O_1 + 2O_2 + 3O_3)}{4}$$

From (2)

$$O_4 = n - (O_1 + O_2 + O_3)$$

Note the equation below does not necessarily hold.

$$\frac{n\hat{\lambda} - (O_1 + 2O_2 + 3O_3)}{4} = n - (O_1 + O_2 + O_3)$$

Assuming  $O_1, O_2$  are free to vary, and let's try to solve  $O_3$  and  $O_4$ :

$$(1) \frac{1}{n} (O_1 + 2O_2 + 3O_3 + 4O_4) = \hat{\lambda}$$

$$(2) O_1 + O_2 + O_3 + O_4 = n$$

Solving this equation group yields the solution of  $O_3$  and  $O_4$

$$O_3 = n(4 - \hat{\lambda}) - 3O_1 - 2O_2$$

$$O_4 = n(\hat{\lambda} - 3) + 2O_1 + O_2$$

This means given random variables  $O_1$  and  $O_2$  and constants related to  $n$  and  $\hat{\lambda}$ , we can solve the equations to find the unique combination of  $O_3$  and  $O_4$  values to make that the two constraints (1) and (2) satisfied. Therefore, the reduction in degrees of freedom is 2.

(Note that it is not guaranteed that in the solution  $O_3$  and  $O_4$  are both positive. This depends on the values of  $O_1, O_2, n, \hat{\lambda}$ .)