Chapter 9-10
Confidence Intervals and Hypothesis Testing
Goodness of Fit Tests

Statistics

Mehmet Güray Güler, PhD

Last updated 28.07.2020

Goodness of Fit Tests

- Goodness of fit test used to determine if a population has a specified theoretical (hypothesized) distribution.
- For that reason we need to have a random sample from the population.
- The test is based on how good a fit we have between
  - the frequency of occurrence of observations in the sample and
  - the expected frequencies obtained from the hypothesized distribution.
- For example, suppose a gambler wishes to see if a given die is balanced.

Goodness of Fit Tests

- EXAMPLE 1. How to test if a given die is balanced.
- We may first roll the die (say) 120 times and count the frequency of occurrence of each face value to obtain a sample distribution as follows:

Face:	**1**	**2**	**3**	**4**	**5**	**6**
Observed	20	22	17	18	19	24
Expected	20	20	20	20	20	20

- Balanced die means all faces have the same probability of occurrence.
- Hence we expect the same frequency for each face.

Goodness of Fit Tests

- H_0: The sample data come from the specified distribution.
- H_1: The sample data do not come from the specified distribution.
- **Test statistic:** $\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$
- **Decision Rule:** Reject H_0 if $\chi^2_{obs} > \chi^2_{\alpha}(k-1)$
- o_i represent the observed frequencies for the ith cell.
- e_i represent the expected frequencies for the ith cell.
- Why should we reject the null hypothesis if the test statistic is larger than the χ^2_{α} value with $(k-1)$ degrees of freedom?

Goodness of Fit Test – Example 1

- EXAMPLE 1. The data for the die rolling experiment:

Face:	1	2	3	4	5	6	
Observed	20	22	17	18	19	24	→ o_i
Expected	20	20	20	20	20	20	→ e_i

- We want to test:
- H_0: The die is balanced (all faces have equal chance)
- H_1: The die is not balanced (some faces appear more frequently than others)

Goodness of Fit Test – Example 1

- Let's calculate the test statistic: .

$$X^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} = \frac{(20-20)^2}{20} + \frac{(22-20)^2}{20} + \cdots + \frac{(24-20)^2}{20} = 1.7$$

- Here we have 6 cells (possible face values), hence the degrees of freedom for the chi-squared statistic is 5. Assume $\alpha = 0.01$.
- **Decision Rule:** Do not reject H_0 since $X^2 = 1.7 < \chi^2_{0.01}(5) = 15.086$.
- P-value $= P(X^2 > 1.7) \approx 0.89$.
- Conclusion: The die seems to be balanced.

Chi-Square Goodness of Fit Test

- EXAMPLE 2. It is claimed that the frequency distribution of the lifetime of a certain brand and model of car battery may be approximated by a normal distribution with the mean $\mu = 3.5$ yrs and the st. Dev. $\sigma = 0.7$ yrs.
- Let's test this by using chi-square goodness of fit test.
- A random sample of 40 batteries were followed and their lifetimes (duration from initial installment time until they fail) were recorded.
- The data are summarized in the frequency table below.

EXAMPLE 2. Battery Lifetime

The lifetime data for the sample of batteries:

Lifetime (yrs)	o_i = Number of Batteries
Class Boundaries	$\boldsymbol{o_i}$
1.45–1.95	2
1.95–2.45	1
2.45–2.95	4
2.95–3.45	15
3.45–3.95	10
3.95–4.45	5
4.45–4.95	3

EXAMPLE 2. Battery Lifetime

- Let's write the hypotheses to be tested:
  - H_0: The data come from $N(\mu = 3.5, \sigma = 0.7)$
  - H_1: No, the data do not come from this distribution.
- The data table shows the observed frequencies for $n = 40$.
- We need to calculate the **expected frequencies** first.
- The test becomes better (more sensitive) when we **combine** the intervals with small frequency so that expected frequency for each cell / class will be **at least 5**

EXAMPLE 2. Battery Lifetime

The lifetime data for the sample of batteries:

Lifetime (yrs)	o_i = Number of Batteries
Class Boundaries	$\boldsymbol{o_i}$
1.45–1.95	2
1.95–2.45	1
2.45–2.95	4
2.95–3.45	15
3.45–3.95	10
3.95–4.45	5
4.45–4.95	3

EXAMPLE 2. Battery Lifetime

- The specified distribution is normal, we will calculate the probability for each interval using EXCEL (or we may look at the table of course)
- Calculate the exp. freq. for the 2nd interval (2.95, 3.45) assuming
  - X is $N(\mu = 3.5, \sigma = 0.7)$.
- From EXCEL we find : $P(2.95<X<3.45) \approx 0.2555$.
- Hence the expected frequency for the second class is:
  - $e_2 = \text{Probability} \times \text{Sample Size} = 0.2555 \times 40 = 10.2204$
- It is customary to round these frequencies to one decimal, hence use **10.2**

EXAMPLE 2. Battery Lifetime

- Now we calculate the probability for 3rd class:
- $P(3.45 < X < 3.95) \approx 0.2683.$
- Hence the expected frequency is: $e_3 = 40 \times 0.2683 \approx 10.7.$
- The first and the last one are calculated somewhat differently.
  - First One: P(X<2.95) = 0.216 ➔ $e_1 = 40 \times 0.216 \approx 8.7$
  - Last One: P(X>3.95) = 0.260 ➔ $e_1 = 40 \times 0.260 \approx 10.4$
- Note that $e_1 + e_2 + e_3 + e_4 = 8.7 + 10.2 + 10.7 + 10.4 = 40.$

EXAMPLE 2. Battery Lifetime

Class Boundaries	o_i		e_i
1.45–1.95	2	7	8.7
1.95–2.45	1		
2.45–2.95	4		
2.95–3.45	15		10.2
3.45–3.95	10		10.7
3.95–4.45	5	8	10.4
4.45–4.95	3		

We now place the expected frequencies in the table and calculate the test statistic using:

$$X^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

EXAMPLE 2. Battery Lifetime

- The degrees of freedom for the chi-squared statistic is one less than the number of terms in the calculation $= 3$.
- Now we compare the computed value of the test statistic with the critical value from the chi-squared distribution with $3\ df$ and the level of significance $\alpha = 0.05$.
- $\chi^2_{obs} = \frac{(7-8.7)^2}{8.7} + \frac{(15-10.2)^2}{10.2} + \frac{(10-10.7)^2}{10.7} + \frac{(8-10.4)^2}{10.4} = 3.15$

EXAMPLE 2. Battery Lifetime

- Decision: Since $\chi^2 = 3.15 < \chi^2_{0.05}(3) = 7.815$,
- We don't reject H_0.
- P-value $= P(\chi^2 > 3.15) \approx 0.37$.
- **Conclusion:** The battery life may have the normal distribution with mean life $\mu = 3.5$ years and the standard deviation $\sigma = 0.7$ years.