

UNIT 5

STATISTICS 2 (SYLLABUS REF. 4.1-4.4)

(Chapter 11, p.316-345)

Aim:

To develop knowledge in statistics and get techniques to analyse and interpret two-variable statistics.

Objective:

LINEAR REGRESSION

- Understand the concept of two-variable statistics (compared to one-variable statistics).
- Understand correlation as the relationship between two numerical variables.
- Understand when this correlation is linear.
- Know what a scatter plot is, and be able to construct one from given data.
- Be able to discover an outlier in a scatter plot.
- Describe the correlation in a scatter plot as *no, weak, moderate, strong, very strong or perfect*, as well as *positive or negative*.
- Interpret the correlation coefficient, r .
- Calculate the correlation coefficient, r , using the feature of calculators.
- Know how to draw the line of best fit, always through the mean and extended to the vertical axis.
- Know how to find the equation of the line of best fit, using the feature of calculators.
- The procedure of using the equation of line of best fit for prediction.
- Know when it is not appropriate to use the equation of the line of best fit for prediction (interpolation vs extrapolation).
- Have an idea of what is meant by "least square regression".
- Work out real life problems including linear regression.

THE CHI-SQUARED TEST OF INDEPENDENCE

- Understand the *contingency table* as a description of the observed values.
- Understand the *table of expected values* as the values we would expect to obtain if the variables were independent.
- The procedure of calculating the values in the table of expected values.

- Be able to use calculator to find the values in the table of expected values.
- Understand that if the difference between the values in the contingency table and the corresponding values in the table of expected values are very small it indicates that the classifications are independent.
- Understand the formula $\chi^2_{calc} = \sum \frac{(f_o - f_e)^2}{f_e}$ and the table that leads to χ^2_{calc} .
- Find the χ^2_{calc} using the feature of calculator.
- Know how to state the *Null hypothesis* (the classifications are independent).
- Know how to state the *Alternative hypothesis* (the classifications are not independent).
- Understand the notation for the hypothesis, as H_0 and H_1 .
- Understand when to *reject* and when to *accept* H_0 and H_1 .
- Know how to find *degrees of freedom*.
- Know that there is a *table of critical values*.
- Be able to find the correct critical value in the table.
- Have an idea of what is meant by *significant level*.
- Solve problems using a Chi-squared test.
- Interpret the p-value given in the calculator, for 5% level as $p < 0.05$ we reject H_0 (the classifications are not independent) and $p > H_0$ we accept H_0 (the classifications are independent).

(The p-value is the probability of obtaining the observed values. If the probability is small it indicates that the classifications are dependent.)

TOK links

When is a statistical test reliable?
How many outliers can there be?

ATL

Discuss projects ideas.
Introduction videos.
Learning checks, quizzes, individual feedback.

Assessment

Formative: Quizzes and homework examination questions.
Summative, included in test exam May 3.