# Statistics and probability

International Baccalaureate · IB Diploma · Mathematics: Analysis and Approaches · HL · Vocabulary in context

| Word | Example |
| --- | --- |
| **population** /ˌpɒpjʊˈleɪʃn/ | Define the **population** before deciding whom to include. |
| **sample** /ˈsæmpl/ | The **sample** contains the people selected for this study. |
| **discrete data** | **Discrete data** can take only certain values, such as the number of children in a family. |
| **continuous data** | **Continuous data** can take any value in a range, such as height or time. |
| **outlier** /ˈaʊtlaɪə/ | Leave out the **outlier** when you draw the line. |
| **mean** /miːn/ | The **mean** is the total divided by the number of values. |
| **median** /ˈmiːdiːən/ | Put the values in order before you find the **median**. |
| **mode** /məʊd/ | The **mode** is the most common value. |
| **quartile** /ˈkwɔːtaɪl/ | The lower **quartile** is a quarter of the way through the data. |
| **interquartile range** /ˌɪntəˈkwɔːtaɪl reɪndʒ/ | Subtract the lower quartile from the upper for the **interquartile range**. |
| **standard deviation** /ˈstændəd ˌdiːvɪˈeɪʃn/ | A larger **standard deviation** means the values are more spread out. |



exercises.lol exercises.com.cn

| Word | Example |
| --- | --- |
| **variance** /'veərɪəns/ | The **variance** is the square of the standard deviation. |
| **cumulative frequency** /'kju:mjʊlətɪv 'fri:kwənsi/ | **Cumulative frequency** adds each count to all earlier counts. |
| **box plot** /bɒks plɒt/ | A **box plot** displays the median and quartiles. |
| **scatter diagram** /'skætə 'daɪəgræm/ | A **scatter diagram** shows two variables as points. |
| **correlation coefficient** /ˌkɒrɪ'leɪʃn ˌkəʊɪ'fɪʃənt/ | The **correlation coefficient** r of a sample estimates the population value $\rho$. $r$ $\rho$ |
| **regression line** /rɪ'greʃn laɪn/ | A **regression line** $y = a + bx$ is the line of best fit for predicting y from x. $y = a + bx$ x y |
| **interpolation** /ɪnˌtɜ:pə'leɪʃn/ | **Interpolation** estimates a value inside the range of the data. |
| **extrapolation** /ek'stræpəleɪʃn/ | Predicting beyond the data is **extrapolation**. |
| **probability** /ˌprɒbə'bɪlɪti/ | A fair coin gives **probability** one half for heads. |
| **sample space** /'sæmpl speɪs/ | The **sample space** is the list of all possible outcomes. |
| **mutually exclusive** /'mju:tʃu:əli eks'klu:sɪv/ | **Mutually exclusive** outcomes cannot occur together. |
| **conditional probability** /kən'dɪʃənl ˌprɒbə'bɪlɪti/ | Find the **conditional probability** using only the chosen group. |
| **independent** /ˌɪndɪ'pendənt/ | **Independent** events do not change each other's probabilities. |



exercises.lol exercises.com.cn

| Word | Example |
| --- | --- |
| **tree diagram** /tri: 'daɪəgræm/ | A **tree diagram** shows each stage of a probability problem. |
| **discrete random variable** /dɪ'skri:t 'rændəm 'veərɪəbl/ | A **discrete random variable** X takes separate values, each with a probability. $X$ |
| **expected value** /ek'spektɪd 'vælju:/ | The **expected value** of one dice roll is 3.5. $3.5$ |
| **normal distribution** /'nɔ:ml ˌdɪstrɪ'bju:ʃn/ | A **normal distribution** has a symmetric bell-shaped density curve. |
| **standardised value** | A **standardised value** of $z = 2$ is two standard deviations above the mean. $z = 2$ |
| **Bayes' theorem** | Use **Bayes' theorem** to update a probability after observing evidence. |
| **continuous random variable** /kən'tɪnju:əs 'rændəm 'veərɪəbl/ | For a **continuous random variable**, probability is an area under a curve. |
| **probability density function** /ˌprɒbə'bɪlɪti 'densɪti 'fʌŋkʃn/ | A **probability density function** describes a continuous random variable. |



exercises.lol exercises.com.cn