Statistics and probability

International Baccalaureate · IB Diploma · Mathematics: Applications and Interpretation · HL · Vocabulary in context

Word	Example
population /ˌpɒpjʊ'leɪʃn/	Define the **population** before deciding whom to include.
sample /'sæmpl/	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 /'aʊtlaɪə/	Leave out the **outlier** when you draw the line.
mean /mi:n/	The **mean** is the total divided by the number of values.
median /'mi:di:ən/	Put the values in order before you find the **median**.
mode /məʊd/	The **mode** is the most common value.
quartile /'kwɔ:taɪl/	The lower **quartile** is a quarter of the way through the data.
interquartile range /ˌɪntə'kwɔ:taɪl reɪndʒ/	Subtract the lower quartile from the upper for the **interquartile range**.
standard deviation /'stændəd ˌdi:vɪ'eɪʃn/	A larger **standard deviation** means the values are more spread out.

exercises.lol exercises.com.cn

Word	Example
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.
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**.
Spearman's rank correlation coefficient /'spɪəmənz ræŋk ˌkɒrɪ'leɪʃn ˌkəʊɪ'fɪʃənt/	**Spearman's rank correlation coefficient** compares two rankings.
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.
independent /ˌɪndɪ'pendənt/	**Independent** events do not change each other's probabilities.
tree diagram /tri: 'daɪəgræm/	A **tree diagram** shows each stage of a probability problem.

Word	Example
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.
null hypothesis /nʌl haɪ'pɒθəsɪs/	The **null hypothesis** specifies the parameter value being tested.
alternative hypothesis /ɔ:l'tɜ:nətɪv haɪ'pɒθəsɪs/	The **alternative hypothesis** says the coin favours heads.
significance level /sɪg'nɪfɪkəns 'levl/	Set the **significance level** before looking at the observed result.
chi-squared test /kaɪ skweəd test/	A **chi-squared test** checks whether our counts are close to the expected counts.
contingency table /kən'tɪndʒənsi 'teɪbl/	A **contingency table** shows counts for two variables.
degrees of freedom /dɪ'gri:z ɒv 'fri:dəm/	A mechanism with one **degree of freedom** needs one motor.
critical value /'krɪtɪkl 'vælju:/	The **critical value** is the edge of the critical region.
goodness of fit test	A **goodness of fit test** checks whether data follow a given distribution.
one-tailed /wʌn teɪld/	A **one-tailed** test checks only whether the mean has gone up.

exercises.lol exercises.com.cn

Word	Example
two-tailed /tu: teɪld/	A **two-tailed** test looks for a difference in either direction.
reliable /rɪ'laɪəbl/	A **reliable** test gives the same result when repeated.
valid	The expansion of $(1 + x)^{-1}$ is **valid** only when x is between -1 and 1. $(1 + x)^{-1}$ x -1 1
residual /rɪ'sɪdʒu:əl/	Subtract the predicted value from the real value to find the **residual**.
coefficient of determination /ˌkəʊɪ'fɪʃənt ɒv dɪˌtɜ:mɪ'neɪʃn/	The **coefficient of determination** r^2 shows how much variation the line explains. r^2
unbiased estimate	The sample mean is an **unbiased estimate** of the population mean.
central limit theorem /'sentrəl 'lɪmɪt 'θɪərəm/	The **central limit theorem** says averages of samples follow a normal curve.
confidence interval /'kɒnfɪdəns 'ɪntəvl/	A narrow **confidence interval** means the estimate is precise.
Poisson distribution /'pɔɪsn ˌdɪstrɪ'bju:ʃn/	The **Poisson distribution** models random events, such as calls per hour.
critical region /'krɪtɪkl 'ri:dʒn/	If the test statistic falls in the **critical region**, we reject the null hypothesis.
Type I error /taɪp aɪ 'erə/	A **Type I error** rejects a true null hypothesis.
Type II error /taɪp 'tu: 'erə/	A **Type II error** fails to reject a false null hypothesis.
Markov chain	In a **Markov chain**, the next state depends only on the current one.

Word	Example
transition matrix /trænˈsɪʃn ˈmeɪtrɪks/	A **transition matrix** shows the chances of moving between states.
steady state /ˈstedi steɪt/	A drug reaches **steady state** after about five half-lives.

exercises.lol exercises.com.cn