

Statistics 1 (S1)

Edexcel IAL Maths · Knowledge organiser

KEY DEFINITIONS

Discrete random variable

Takes separate values, each with a probability; the probabilities add to 1.

Expected value

$E(X) = \sum x P(X = x)$, the long-run mean.

Product moment correlation coefficient

$r = S_{xy}/\sqrt{(S_{xx} S_{yy})}$ measures linear correlation, from -1 to 1 .

Interpolation

Estimating the median or a quartile from grouped data, assuming values are spread evenly in each class.

YOU SHOULD BE ABLE TO...

- Explain how statistical models are used and refined, and classify data as qualitative or quantitative and discrete or continuous.
- Draw and interpret stem-and-leaf diagrams, including back-to-back diagrams, and identify outliers using a rule based on the quartiles.
- Use the rules for mutually exclusive and independent events, and test whether two events are independent using $P(A \cap B) = P(A)P(B)$.
- Calculate S_{xx} , S_{yy} and S_{xy} from summary statistics and use them to find the product moment correlation coefficient, interpreting its value.
- Calculate the expectation $E(X)$ and the variance $\text{Var}(X)$ of a discrete random variable from its probability distribution.
- Solve normal distribution problems in context, such as the proportion of items rejected, combining normal probabilities with independent events.

COMMON MISTAKES

- ✗ Standardisation not shown when the question requires it
- ✗ Outliers: rule applied but no conclusion
- ✗ Interpolation using class midpoints
- ✗ Variance formula missing the $-\bar{x}^2$

KEY FORMULAS

B = given in the exam (formula booklet or list); learn the rest.

Mean; variance $\bar{x} = \frac{\sum fx}{\sum f}$, $\sigma^2 = \frac{\sum fx^2}{\sum f} - \bar{x}^2$

Coding $y = \frac{x-a}{b}$ $\bar{x} = a + b\bar{y}$, $\sigma_x = |b|\sigma_y$

Median by interpolation $Q_2 = L + \frac{\frac{n}{2} - F}{f} \times w$

Outliers (if asked) $\backslash$ (or $> Q_3 + k(Q_3 - Q_1)$

Union **B** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Intersection **B** $P(A \cap B) = P(A)P(B | A)$

Conditional; independent $P(A | B) = \frac{P(A \cap B)}{P(B)}$; $P(A \cap B) = P(A)P(B)$

Discrete r.v. **B** $E(X) = \sum xP(X = x)$,
 $\text{Var}(X) = \sum x^2P(X = x) - \mu^2$

Function of X **B** $E(g(X)) = \sum g(x)P(X = x)$

Linear $E(aX + b) = aE(X) + b$,
 $\text{Var}(aX + b) = a^2\text{Var}(X)$

Discrete uniform on $1, \dots, n$ $E(X) = \frac{n+1}{2}$, $\text{Var}(X) = \frac{n^2-1}{12}$

S_{xx} , S_{xy} **B** $S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$, $S_{xy} = \sum xy - \frac{\sum x \sum y}{n}$

PMCC **B** $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$

Regression line $y = a + bx$ **B** $b = \frac{S_{xy}}{S_{xx}}$, $a = \bar{y} - b\bar{x}$

Normal $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

WORKED EXAMPLE

X takes the values 1, 2 and 3 with probabilities 0.2, 0.5 and 0.3. Find $E(X)$ and $\text{Var}(X)$.

1. $E(X) = 1(0.2) + 2(0.5) + 3(0.3) = 2.1$

2. $E(X^2) = 1(0.2) + 4(0.5) + 9(0.3) = 4.9$

3. $\text{Var}(X) = 4.9 - 2.1^2$

Answer: $E(X) = 2.1$ and $\text{Var}(X) = 0.49$