

Statistics 1

Edexcel IAL S1 (WST01/01); a permitted applied unit of Further Mathematics
after the pure units

1. Unit S1 at a glance

Unit S1 at a glance

WST01/01: 1 hour 30 minutes, 75 marks; P1–P2 assumed. Must-know (not in the booklet): $\bar{x} = \frac{\sum x}{n}$, the addition and multiplication laws, $E(aX + b) = aE(X) + b$, $\text{Var}(aX + b) = a^2 \text{Var}(X)$, and $Z = \frac{X - \mu}{\sigma}$.

2. Mathematical models

Why models

A statistical model simplifies a real situation to represent, analyse and predict it: quick, cheap, repeatable, refinable. Quote a reason BEYOND prediction for the mark.

3. Representing data

Histograms and box plots

Histogram: height = f/w , area = frequency — compare bars by RATIO of densities. WST01's tallest bar 10 cm makes the second-tallest $\frac{10.8}{18} \times 10 = 6$ cm. Box plots: quartiles and whiskers, outliers as crosses under the stated rule.

4. Location and dispersion

Combining and coding

Merge groups by totals: $\sum x = n_1 \bar{x}_1 + n_2 \bar{x}_2$, $\sum x^2$ adds, then $\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2$.

Interpolate quartiles inside the containing class. Coding reverses: $w = 11.5d$ gives

$$S_{tw} = 11.5 S_{td}, S_{ww} = 11.5^2 S_{dd}.$$

5. Skewness

Skewness

Compare $(Q_3 - Q_2)$ with $(Q_2 - Q_1)$, or mean with median. WST01's speeds: $5.875 > 3.11$, and mean $30.6 >$ median 28.25 : positively skewed, so a Normal model is refused.

6. Probability

Probability

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$; $P(A \cap B) = P(A)P(B|A)$; reversal
 $P(B|A) = \frac{P(A \cap B)}{P(A)}$. Trees: without replacement drops both parts by one. WST01:
 $0.2 \times 0.45 = 0.09$ then $P(C) = 0.23$; beads $1/2$, $7/12$, $1/7$.

7. Correlation and regression

Regression

$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$; y on x : $b = \frac{S_{xy}}{S_{xx}}$, $a = \bar{y} - b\bar{x}$. WST01 coded data: $r = -0.844$,
 $w = -3.46t + 30.9$, so $d = -0.301t + 2.69$; predict only inside the range.

8. Discrete random variables

Discrete DRVs

Solve inequalities for X first; $E(X) = \sum xp(x)$, $E(X^2) = \sum x^2p(x)$,
 $\text{Var} = E(X^2) - [E(X)]^2$. WST01: $F(3.5) = 0.85$, $P(3X - 3 < X + 2) = 0.5$, $E = 2.35$,
 $\text{Var} = 1.1275$, $\text{Var}(5 - 2X) = 4.51$. Uniform on $1..n$: mean $\frac{n+1}{2}$, variance $\frac{n^2-1}{12}$.

9. The Normal distribution

Normal

$Z = \frac{X - \mu}{\sigma}$, both directions; symmetry pairs opposite tails. WST01 bags:

$P(X > 508) = P(Z > 2) = 0.0228$, symmetric band 0.9544, exactly 2 of 4 gives

$6p^2(1 - p)^2 = 0.00298$; σ from a percentile; equal-tail conditions give simultaneous equations, $k = 505$.

Check yourself

Check yourself

Reasons for models beyond prediction; density ratios and interpolation quartiles; merging via totals; coding effects on mean, variance, SD; Venn and tree probability with conditional reversal; r from S statistics with in-context reading; the discrete DRV routine; Normal standardising both ways with symmetry and simultaneous finishes.