

Statistics 3

Edexcel IAL S3 (WST03/01); the final statistics unit completing the further-mathematics table

1. Unit S3 at a glance

Unit S3 at a glance

WST03/01: 1 hour 30 minutes, 75 marks; the booklet carries the Normal and chi-squared tables plus Spearman critical values. Must-know:

$$aX \pm bY \sim N(a\mu_x \pm b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2); \bar{X} \sim N(\mu, \sigma^2/n); s^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1};$$

$$\chi^2 = \sum \frac{(O-E)^2}{E}; r_s = 1 - \frac{6 \sum d^2}{n(n^2-1)}.$$

2. Combinations of random variables

Combinations

Independent Normals combine: add means, add squared coefficients times variances. $S_1 + S_2 + S_3$ with $S \sim N(18, 0.03^2)$ is $N(54, 0.0027)$, while $3S$ has nine times the variance; $C_1 - C_2$ is $N(0, 2\sigma^2)$ with $P(|\cdot| > c)$ doubled. WME03's Q8c: 28 sand + 5 cement - 29 pallets is $N(194, 33.6697)$ — the 29 squares.

3. Sampling methods

Sampling methods

Simple random needs a frame; systematic takes every N/n -th from a random start (spacing locks out neighbours: members 5 and 6 cannot co-occur); stratified samples strata proportionally; quota fills interviewer counts with no frame. Give the mechanical reason, not a generic speed claim.

4. Estimators and the sample mean

Estimators

A statistic has no unknown parameters; bias is $E(\hat{\theta}) - \theta$. $\bar{X}$ and s^2 (with $n - 1$) are unbiased. Two probability statements about $\bar{X}$ invert to two z-equations: WST03 Q7's 0.3446 and 0.0179 give $z = -0.4$ and 2.1 , hence $\sigma = 12\sqrt{n}$, $\mu = 50$.

5. Confidence intervals and tests

Intervals and tests

CI: $\bar{x} \pm z\sigma/\sqrt{n}$; two samples with a gap: $z = \frac{(\bar{x}-\bar{y})-(\mu_x-\mu_y)}{\sqrt{\sigma_x^2/n_x+\sigma_y^2/n_y}}$. WST03 Q4: grey vs red squirrels with H_0 gap 200 g gives $z = 2.457 > 1.6449$ — the gap is exceeded. Large samples: substitute s^2 (CLT licenses non-Normal populations).

6. Chi-squared

Chi-squared

Fit: $E = np_i$ from the model, $df = \text{cells} - 1 - (\text{estimated parameters})$: the litter data with $\hat{p} = 0.488$ gives $X^2 = 10.95$ on 4 df — binomial rejected. Contingency: $E = \frac{\text{row} \times \text{col}}{\text{total}}$, $df = (r - 1)(c - 1)$: the engine table gives $X^2 = 12.325 > 9.488$ — association. Combine cells with $E < 5$ first.

7. Rank correlation

Rank correlation

$r_s = 1 - \frac{6 \sum d^2}{n(n^2-1)}$: the judges' ranks with $\sum d^2 = 66$ give $r_s = 0.6$. Test $H_0: \rho = 0$ against the table: the 1% one-tailed critical value 0.7455 is not reached — insufficient evidence of positive agreement.

Check yourself

Check yourself

Combine independent Normals (differences and absolute differences); four sampling methods with mechanical reasons; bias algebra and paired unbiased-estimator equations; recovering (μ, σ) from two sample-mean probabilities; z intervals and one/two-sample tests with the variance assumption; chi-squared for fit and contingency with df bookkeeping; Spearman's r_s and its table test.